



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

DIPLOMOVÁ PRÁCE

Bc. Jakub Filka

Projekce úmrtnostních tabulek a jejich vplyv na implicitní hodnotu pojišťovny

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Michal Pešta, PhD.

Studijní program: Matematika

Studijní obor: Finanční a pojistná matematika

Praha 2016

Chcel by som sa touto cestou poďakovať vedúcemu diplomovej práce, RNDr. Michalovi Peštovi, Ph.D., za všetky matematické a technické postrehy k práci a za jeho príkladnú ochotu a dobrú vôľu pomôcť študentovi pri akýchkoľvek problémoch a otázkach, Mgr. Jaroslavovi Dufekovi, za odborné poradenstvo v oblasti úmrtnostných tabuliek, MCEV a za pomoc pri výbere diplomovej práce, RNDr. Borisovi Burcinovi, Ph.D. za množstvo odborných konzultácií v oblasti demografie a za poskytnutie softwaru DERAS pre účely tejto práce, RNDr. Lucii Mazurovej, Ph.D., za vedomosti nadobudnuté na jej hodinách demografie a účtovníctva poisťovní, ktoré mi pomohli pri písaní tejto práce.

Prohlašuji, že jsem tuto diplomovou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Název práce: Projekce úmrtnostních tabulek a jejich vplyv na implicitní hodnotu pojišťovny

Autor: Jakub Filka

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Michal Pešta, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: V práci se věnujeme vývoji úmrtnostních tabulek v České republice počínaje rokem 1950. Naším cílem je prostudovat 6 základních modelů, které mohou být použité pro modelování úmrtnosti lidí starších 60 let. Mezi zkoumanými modely jsou model Gompertz-Makehama, logistické modely Thatchera a Kannista, modely Coale-Kiskera a Heligman-Pollarda. Naše analýza se soustředí na projekční vlastnosti modelů do let nejvyšších. Ukázalo se, že trend úmrtnosti se nejlíp daří zachytit logistickým modelům, které na rozdíl od Gompertz-Makehamova modelu nenadhodnocují pravděpodobnosti umírání v letech nejvyšších, a to především pro ženy, kde data nevykazují takovou disperzi jako u mužů.

Klíčová slova: projekce úmrtnostních tabulek, Gompertz-Makeham, logistické modely

Title: Projection of mortality tables and its influence on embedded value of insurance

Author: Jakub Filka

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Pešta Michal, Dr.,

Department: Department of Probability and Mathematical Statistics

Abstract: We study development of mortality tables from 1950 to present in Czech Republic. Our aim is to look at the 6 basic models, which can be potentially used to describe behavior of dying for people over 60 years. Models that are being investigated vary from generally accepted Gompertz-Makeham model to logistic models of Thatcher and Kannisto. We also introduce Coale-Kisker and Heligman-Pollard model. Our analysis is concentrated mostly on projecting abilities of given models to the highest ages. Especially for women, where data do not show such dispersion as in the case of men, there is a visible trend that can be described better by using logistic models instead of Gompertz-Makeham model, which has a tendency to overestimate the probabilities of dying in higher ages.

Keywords: projection of mortality tables, Gompertz-Makeham, logistic models

Názov práce: Projekcia úmrtnostných tabuliek a ich vplyv na implicitnú hodnotu poisťovne.

Autor: Jakub Filka

Katedra: Katedra pravdepodobnosti a matematickej statistiky

Vedúci diplomovej práce: RNDr. Pešta Michal, Ph.D., Katedra pravdepodobnosti a matematickej statistiky

Abstrakt: V práci sa venujeme vývoju úmrtnostných tabuliek v Českej republike, počínajúc rokom 1950. Naším cieľom je preštudovať 6 základných modelov, ktoré môžu byť použité pre modelovanie úmrtnosti vo veku 60 a viac. Medzi skúmanými modelmi sú modely Gompertz-Makehama, logistické modely Thatcher a Kannista, modely Coale-Kiskera a Heligman-Pollarda. Naša analýza sa sústreďuje na projekčné vlastnosti modelov do najvyššieho veku. Najmä pre ženy, kde dáta nevykazujú takú disperziu ako u mužov, sa ukázalo, že trend úmrtnosti sa najlepšie darí zachytiť logistickým modelom, ktoré na rozdiel od Gompertz-Makehama nenahodnocujú pravdepodobnosti úmrtnosti vo vysokých vekoch.

Kľúčové slová: projekcia úmrtnostných tabuliek, Gompertz-Makeham, logistické modely

Obsah

1	Úvod	2
2	Úmrtnostné tabuľky	3
2.1	Motivácia a delenie úmrtnostných tabuliek	3
2.2	Popis základných charakteristík	4
2.3	Porovnanie metodológií	6
3	Modely projekcie úmrtnosti	7
3.1	Gompertz-Makehamov model	7
3.2	Coale-Kiskerov model	8
3.3	Kannisto a Thatcherov model	9
3.4	Helligman-Pollardov model	10
3.5	Odhady parametrov	11
3.5.1	Gauss-Newtonov algoritmus	11
3.5.2	Levenberg-Marquadtové metódy	13
3.6	Intervaly spoľahlivosti	14
3.6.1	Predikčné intervaly pre rastovú krivku	14
4	Analýza úmrtnosti	16
4.1	Úmrtnosť v rokoch 1950 až 1999	16
4.2	Vývoj úmrtnosti v posledných 15 rokoch	18
4.3	Vhodnosť jednotlivých modelov na populáciu Českej republiky . .	19
4.4	Predikčné intervaly pre projektované veky	23
5	Vplyv na implicitnú hodnotu poisťovne	26
5.1	Implicitná hodnota poisťovne	26
5.2	Dopad na poisťovňu	28
6	Záver	31
	Literatúra	33
	Zoznam obrázkov	35

Kapitola 1

Úvod

Tvorba úmrtnostných tabuliek je dôležitou súčasťou životného a penzijného poistenia. Od začiatku 21. storočia v Českej republike významne narastá očakávaná dĺžka života, ktorá súvisí so stále väčším počtom ľudí dožívajúcich sa vysokého veku. To nám poskytuje dostatočné množstvo dát pre posúdenie kvality potenciálnych modelov, ktoré sa snažia proces umierania zachytiť.

Cieľom tejto práce je predstaviť hlavné modely, ktoré sa používajú pri tvorbe úmrtnostných tabuliek pre projekciu úmrtnosti do vysokých vekov a analyzovať ich vhodnosť pre českú populáciu.

V teoretickej časti sa venujeme predstaveniu jednotlivých modelov. Na jednej strane máme oficiálne používaný Gompertz-Makehamov a Coale-Kiskerov model, ktoré boli v minulosti považované za najadekvátnejšie aj v ostatných štátoch Európy. Avšak momentálne sú tieto modely pod paľbou ostrej kritiky vzhľadom na preukázané nadhodnocovanie pravdepodobností úmrtia. Na strane druhej sú logistické modely, ktoré sa v posledných rokoch tešia stále väčšej a väčšej obľube a v najvyspelejších štátoch Európy sú uznávané ako oficiálne modely pre tvorbu úmrtnostných tabuliek vo vysokých vekoch. Okrem nich si v práci predstavíme aj model Heligman-Pollarda, ktorý môže poslúžiť ako kompromis medzi týmito dvoma prístupmi. Našou základnou úlohou je odhadnúť parametre vybraných modelov a získať intervaly spoľahlivosti pre skúmané vekové kategórie.

V praktickej časti našej práce sa následne venujeme konštrukcii úmrtnostných tabuliek, ktoré v rámci vybraných modelov porovnávame so skutočnosťou. Začíname rokom 1950 a postupujeme až do súčasnosti, zvlášť pre mužov a zvlášť pre ženy. Hlavným cieľom je porovnanie projekčných vlastností jednotlivých modelov pre dáta Českej republiky. Využijeme takisto projekt Human Mortality Database.

V poslednej kapitole si predstavíme jeden zo základných ukazovateľov poisťovne — pojem implicitná hodnota. Pomocou rozkladu očakávaného zisku, ale aj na základe jednoduchého príkladu, kde si vysvetlíme, aký významný dopad majú úmrtnostné tabuľky na tento ukazovateľ.

Kapitola 2

Úmrtnostné tabuľky

V úvodnej kapitole sa budeme zaoberať všeobecným popisom úmrtnostných tabuliek a motiváciou ich tvorby v súvislosti so životným poistením. Zavedieme základné delenie úmrtnostných tabuliek a popíšeme charakteristiky potrebné pre ich konštrukciu v Českej republike. Nakoniec porovnáme oficiálnu metodológiu konštrukcie, ktorú využíva Český štatistický úrad s postupom, ktorý budeme v tejto práci používať. Budeme pritom vychádzať z publikácií [3] a [13].

2.1 Motivácia a delenie úmrtnostných tabuliek

Pri pohľade do histórie úmrtnostných tabuliek môžeme objaviť, že 1. úplná úmrtnostná tabuľka pochádza z roku 1693, ktorý je označovaný aj ako rok začiatku poistenia. Zostavil ju E. Halley na základe matričných záznamov Vratislavy za roky 1687 až 1691. Tvorba úmrtnostných tabuliek je už od samého začiatku úzko spojená so životným poistením. Motiváciu jej tvorby vysvetlíme na príklade poistenia v prípade smrti, ktoré pracuje na nasledujúcom princípe. Pokiaľ dôjde k úmrtiu poistenej osoby, vyplatí poisťovňa poistné plnenie niekomu z pozostalých, respektíve osobe zjednanej v poistnej zmluve. Pri uzavretí tohto druhu poistenia je poisťovňa kompenzovaná vystavenému riziku prostredníctvom prijatého poistného. Pre stanovenie jeho výšky je dôležité odhadnúť pravdepodobnosť, s akou bude poisťovňa poistné plnenie vyplácať. Tento odhad vychádza najmä zo sledovania pravdepodobnosti úmrtia v danom štáte, ktorá býva zachytená práve v úmrtnostných tabuľkách.

Základné delenie úmrtnostných tabuliek sa člení na tabuľky úplné, ktoré obsahujú každý celočíselný vek, a skrátené, ktoré sa obvykle vyznačujú 5 ročnými vekovými intervalmi. Okrem toho môžeme úmrtnostné tabuľky rozdeliť na prierezové (bežné), tie vychádzajú z údajov o populácii za krátke časové obdobie (napr. kalendárny rok), a kohortné, kde sa sleduje úmrtnosť celej generácie — kohorty. Tieto tabuľky sa využívajú napr. pri sledovaní umierania pacientov na nejakú chorobu. Existujú aj mnohé ďalšie delenia úmrtnostných tabuliek, pre účely tejto práce je ale toto delenie postačujúce. Budeme sa zaoberať úplnými a prierezovými tabuľkami, pričom základnou periódou bude pre nás jeden kalendárny rok.

2.2 Popis základných charakteristík

Ako sme už spomínali vyššie, základom pre nás budú pravdepodobnosti úmrtia, pričom pravdepodobnosť úmrtia jedinca vo veku x , to jest pravdepodobnosť, že sa jedinec nedožije veku $x+1$ označíme q_x . Pravdepodobnosť, že sa jedinec vo veku x dožije veku $x+1$ označíme p_x . Zrejme teda platí vzťah

$$q_x + p_x = 1. \quad (2.1)$$

Výpočet vekovo špecifických mier úmrtnosti pre daný kalendárny rok vychádza zo skutočného počtu mŕtvych obyvateľov D_x vo veku x a z počtu obyvateľov vo veku x k 1. 7. daného roku, takzvaný stredný stav populácie P_x .

Pravdepodobnosti úmrtia vo veku x sa počítajú zo vzťahu

$$q_x = 1 - e^{-m_x}, \quad (2.2)$$

kde m_x značí pomer celkového počtu zomrelých jedincov D_x vo veku x a vekovo príslušného stredného stavu populácie P_x . Tento pomer nazývame špecifická miera úmrtnosti:

$$m_x = \frac{D_x}{P_x}. \quad (2.3)$$

Výnimku tvorí len vek $x = 0$, pre ktorý sa pravdepodobnosť úmrtia (tzv. kojenecká úmrtnosť) spočíta zo vzťahu

$$q_0 = \frac{D_0}{N}, \quad (2.4)$$

kde N značí počet úspešne novonarodených osôb v danom období. Táto výnimka je spôsobená vyššou úmrtnosťou v prvých týždňoch života dieťaťa, čo má za následok, že úmrtnosť nie je rovnomerne rozložená v rámci nultého veku.

Pre účely praktických výpočtov sa často používa aproximácia

$$\begin{aligned} q_x &= 1 - e^{-m_x} = 1 - \left(\sum_{n=0}^{\infty} \frac{(-m_x)^n}{n!} \right) \\ &= 1 - \left(1 - m_x + \frac{m_x^2}{2} - \frac{m_x^3}{6} + \dots \right) \approx m_x - \frac{m_x^2}{2} + \frac{m_x^3}{6}, \end{aligned} \quad (2.5)$$

ktorú dostaneme s využitím Taylorovho rozvoja exponenciály. Špecifické miery úmrtnosti sa v najvyššom veku pohybujú okolo hodnoty 0.5, pre ktorú je táto aproximácia uspokojivo presná a v praktickej časti ju budeme používať na zachytenie skutočnosti.

Ďalším ukazovateľom, ktorý môžeme nájsť v úmrtnostnej tabuľke, je tabuľkový počet ľudí dožívajúcich sa veku x . Ide o hypotetický počet osôb, ktorý sa dožije veku x zo 100 000 úspešne narodených v prípade, že sa priebeh úmrtnosti počas života týchto ľudí nezmení. Tento ukazovateľ sa označuje symbolom l_x a napočítava sa rekurentným vzťahom

$$l_{x+1} = l_x p_x, \quad (2.6)$$

pričom $l_0 = 100\,000$ (koreň úmrtnostnej tabuľky).

Podobne tabuľkový počet mŕtvych (hypotetický počet mŕtvych vo veku x) budeme značiť symbolom d_x a počítať zo vzťahu

$$d_x = l_x - l_{x+1}. \quad (2.7)$$

Tabuľkový počet žijúcich (hypotetický priemerný počet ľudí žijúcich vo veku x), ktorý si môžeme predstaviť aj ako celkovú dobu prežitú tabuľkovou generáciou vo veku x , budeme značiť L_x a počítať zo vzťahu

$$L_x = l_{x+1} + a_x d_x, \quad (2.8)$$

kde a_x predstavuje priemernú dobu prežitú vo veku x jedincom, ktorý v tomto roku umrie. V prípade, že predpokladáme, že intenzita úmrtnosti sa v priebehu roku výrazne nemení, kladieme $a_x = 1/2$ pre $x = 1, 2, \dots$ (pre $x = 0$ všeobecne neplatí, pozri kojeneckú úmrtnosť). Dostávame tak vzťah

$$L_x = \frac{1}{2}(l_{x+1} + l_x). \quad (2.9)$$

Ďalším ukazovateľom je pomocný ukazovateľ T_x , ktorý udáva celkovú dobu prežitia tabuľkovou generáciou vo veku vyššom, než je vek x . Dostaneme ho zo vzťahu

$$T_x = L_x + L_{x+1} + L_{x+2} + \dots + L_\omega, \quad (2.10)$$

kde ω značí najvyšší vek v tabuľke úmrtnosti (pre Českú republiku momentálne stanovený na $\omega = 105$). Spolu s celkovou dobou prežitou tabuľkovou generáciou vo veku x (L_x) slúži na výpočet strednej dĺžky života e_x , tzv. nádeje dožitia pri zachovaní priebehu úmrtnosti. Počítame zo vzťahu:

$$e_x = \frac{T_x}{L_x}. \quad (2.11)$$

Stredná dĺžka života teda hovorí o priemernej dobe, ktorú má nádej prežiť x -ročná osoba pri zachovaní úmrtnosti sledovaného obdobia. Ide o syntetický ukazovateľ, ktorý odráža pomery vo všetkých vekových skupinách.

Pre účely matematického modelovania je dôležité zaviesť pojem *intenzity úmrtnosti*, ktorý budeme označovať symbolom μ_x . Uvažujme model náhodnej dĺžky života. Nech X je náhodná veličina predstavujúca vek, ktorý sa dožije novonarodený jedinec. Pravdepodobnosť, že osoba vo veku x sa nedožije veku $x + t$ budeme značiť symbolom ${}_x q_t$ a pravdepodobnosť, že osoba vo veku x sa dožije veku $x + t$ symbolom ${}_x p_t$. Potom $P(X > x)$ je ${}_x p_0$ (pravdepodobnosť prežitia) a platí

$$\begin{aligned} \mu_x &= \lim_{h \rightarrow 0} \frac{P(x < X < x + h | X > x)}{h} = \\ &= \lim_{h \rightarrow 0} \frac{P(X > x) - P(X > x + h)}{h P(X > x)} = -\frac{\partial_x p_0}{p_0} \frac{1}{\partial_x}. \end{aligned} \quad (2.12)$$

2.3 Porovnanie metodológií

Český statistický úrad pri konštrukcii úmrtnostných tabuliek postupuje nasledovne. K výpočtom využíva údaje o celkovom počte mŕtvych jedincov D_x vo veku x , a o vekovo príslušnom strednom stave populácie P_x , spolu so živo narodenými jedincami.

Kojeneckú úmrtnosť určí podľa vzťahu (2.4). Veky 1 — 3 sú potom určené na základe špecifických mier úmrtností. Od veku 4 až do veku y , ktorý spravidla nadobúda hodnoty okolo 70, vyrovná pravdepodobnosti úmrtia podľa vzorca

$$q_x^{vyrovn} = [105q_x + 90(q_{x-1} + q_{x+1}) + 45(q_{x-2} + q_{x+2}) - 30(q_{x-3} + q_{x+3})]/315. \quad (2.13)$$

Tým sa odstránia náhodné výkyvy a krivka sa vyrovná. Roky $y+$, to jest spravidla roky 71 až 105, modeluje pomocou Gompertz-Makehamovho modelu. Odhady parametrov sú napočítané tzv. King-Hardyho metódou. Spočítajú sa modelové pravdepodobnosti dožitia

$$r_x = e^{a+b \cdot x} \quad (2.14)$$

a nájde sa vek y ($y \geq 75$), pre ktorý nadobúda odchýlka

$$|p_x^{vyrovn} - r_x| \quad (2.15)$$

minimálne hodnoty. Od veku y potom pravdepodobnosti úmrtia popisujeme krivkou $q_x^{GM}(1 - r_x)$. Prechod na extrapolované hodnoty od veku približne 80+, ktorý je opodstatnený vzhľadom na malú veľkosť súboru mŕtvych, je upravený vyrovnaním hodnôt pre roky $x = (y - 4), \dots, (y + 4)$

$$q_x^{GM} = 1 - \left[\left(1 - \frac{x - y + 5}{10} \right) p_x^{vyrovn} + \frac{x - y + 5}{10} r_x \right]. \quad (2.16)$$

V našej práci budeme postupovať veľmi podobným spôsobom a zameriame sa práve na projekciu najvyšších rokov. Vstupné dáta budú úplne totožné. Modely založíme na údajoch o úmrtnosti v rokoch 60 až 85, ktoré následne extrapolujeme až do maximálneho možného veku ω . Relatívne malá spodná veková hranica je volená najmä z požiadavky nájdenia vhodného modelu. Horná veková hranica 85 rokov, na určenie regresného modelu, je väčšia než tá, ktorú používa ČSÚ. Pri jej voľbe vychádzame z presvedčenia expertov, že horná veková hranica pre spoľahlivé dáta v posledných rokoch narástla, a to následne potvrdíme analýzou úmrtnosti v úvode tretej kapitoly. Po matematickej stránke to pre nás znamená viac pozorovaní, čo môže presnosti modelu len napomôcť. V práci sa budeme zaoberať aj vhodnou voľbou maximálneho veku ω pre českú populáciu. Pozrieme sa na roky 105, 110 a na prvý pohľad trochu optimistické roky 115 a 120, ktoré ale môžu byť prínosom pre čitateľov v budúcnosti.

Predstavíme si 6 najrozšírenejších demografických modelov pre úmrtnosti vo vysokom veku a pozrieme sa na ich vyrovnanie a projekciu pre najvyššie roky. Dopodrobna tak zanalyzujeme oficiálnu metodológiu Českého statistického úradu a iné možné postupy, ktoré sa používajú v najrozvinutejších štátoch sveta.

Kapitola 3

Modely projekcie úmrtnosti

V tejto kapitole si predstavíme vybrané stochastické modely projekcie úmrtnosti s teoretickým základom potrebným pre praktickú analýzu dopadov jednotlivých metód na ukazovatele poisťovne. Krivku pravdepodobností úmrtia skrz všetky roky možno podľa [7] rozdeliť na 4 časti, kde prvá časť modeluje kojeneckú úmrtnosť, druhá časť úmrtnosť do 20 rokov, tretia časť takzvanú normálnu, pracujúcu úmrtnosť a štvrtá časť sa zaoberá úmrtnosťou vo vyšších rokoch. V našej práci sa zameriame na práve túto poslednú časť a predstavíme si modely, ktoré sú považované za najadekvátnejšie vo svete demografie. Zameriame sa, tým pádom najmä na riziko dlhovekosti, ktoré predstavuje najväčiu hrozbu pri takzvaných annuitných produktoch, a budeme skúmať ako sa s ním jednotlivé modely vysporiadajú.

Základnou charakteristikou stochastického prístupu je fakt, že vek jedinca x je považovaný za náhodnú premennú, teda má zmysel zaoberať sa jeho pravdepodobnostným rozdelením. Tento aparát nám poskytuje výhodu oproti ostatným prístupom v podobe získania nielen bodových, ale aj intervalových odhadov skúmaných charakteristík ako sú pravdepodobnosti úmrtia a iné. Pri predstavovaní modelov budeme vo všeobecnosti vychádzať z publikácie [1].

3.1 Gompertz-Makehamov model

Dlhú dobu bol za najlepší model považovaný Gompertz-Makehamov model, ktorý predpokladá, že intenzita úmrtnosti sa mení s vekom exponenciálne. Základnú štruktúru modelu položil v roku 1825 Gompertz, ktorý vychádzal z faktu, že schopnosť jedinca odolávať deštrukcii klesá s vekom úmerne veľkosti tejto schopnosti. Ak chápeme schopnosť jedinca odolávať deštrukcii ako prevrátenú hodnotu intenzity úmrtnosti, môžeme toto tvrdenie prepísať

$$\frac{d\left(\frac{1}{\mu_x}\right)}{dx} = -h \frac{1}{\mu_x}, \quad h > 0. \quad (3.1)$$

Integráciou oboch strán a vyjadrením μ_x dostávame

$$\log\left(\frac{1}{\mu_x}\right) = -h x + c, \quad h > 0, \quad (3.2)$$

kde c je konštanta. Úpravou a položením $b = e^{-c}$, $c = e^h$ dostávame tvar Gompertzovho modelu

$$\mu_x = b c^x, \quad (3.3)$$

pre $0 < b < 1$, $c > 1$, ktorý nám vyjadruje exponenciálnu závislosť úmrtnosti na veku jedinca prostredníctvom parametrov b a c . V roku 1887 prišiel Makeham s myšlienkou, že úmrtnosť závisí aj od iných faktoroch, než len od veku jedinca, čím vznikol výsledný model

$$\mu_x = a + b c^x, \quad (3.4)$$

kde opäť $0 < b < 1$, $c > 1$ a parameter a zachytáva ostatné, na veku nezávislé príčiny úmrtnosti (úrazy, nehody a pod.).

Výhodou tejto metódy je jednoduchá interpretácia, kompatibilita, spojená s relatívne malým množstvom parametrov a tiež fakt, že k odhadu parametrov modelu nie je potrebný žiadny špeciálny software. Na druhej strane je ale táto metóda pesimistická v úmrtnosti vysokých rokov, ktorú značne nadhodnocuje. Intenzita úmrtnosti narastá exponenciálne s vekom, a keďže

$$p_x = \exp\left(-\int_0^1 \mu_{x+t}\right) \sim \exp\left(-\mu_{x+\frac{1}{2}}\right), \quad (3.5)$$

z Taylorovej aproximácie v strede intervalu, t. j. ak pravdepodobnosť prežitia klesá limitne ku 0, pravdepodobnosť úmrtia sa blíži k 1 pre najvyššiu vekovú kategóriu, čo odporuje trendu vývoja úmrtnosti pre danú vekovú kategóriu v posledných rokoch. Napriek tomu je tento model považovaný za najadekvátnejší.

Zaujímavé je, že Gompertz a Makeham si uvedomovali, že vzťah (3.1) rozhodne nezodpovedá súhrnnému vekovému rozpätiu [12]. Tento vzťah bol Gompertzom pôvodne určený iba pre vekové hranice 20 až 60 rokov [7], avšak v súčasnej dobe sa používa práve pre roky najvyššie. To je aj hlavným dôvodom ustavičných podnetov na otázky vynárajúce sa v popredných kruhoch demografie, keďže limity Gompertzovho zákona boli mnohými autormi potvrdené. Mildvan a Strehler vo svojej publikácii ostro kritizovali popísaný zákon tvrdením, že pravdepodobnosti úmrtia konvergujú v rokoch najvyšších progresívne pomalšie ako je tomu u tohto exponenciálneho modelu [11]. Podobnú kritiku možno nájsť u mnohých ďalších autorov: [6],[8],[9]. Vhodnosť tohto modelu môže byť vysvetlená vysokým nárastom životnej úrovne v poslednom storočí, čo mohlo určenú vekovú hranicu posunúť vyššie.

3.2 Coale-Kiskerov model

Jednou z alternatív pre Gompertz-Makehamov model je model Coale-Kiskera, ktorý kladie dôraz na zmenu miery úmrtnosti 2 v po sebe nasledujúcich rokoch. Predpokladá, že nárast miery úmrtnosti v najvyšších rokoch klesá lineárne nárastom od Gompertz-Makehamovho modelu, kde predpokladáme konštantný priebeh. Na tomto predpoklade bola založená premenná k_x tvaru

$$k_x = \log\left(\frac{m_x}{m_{x+1}}\right), \quad (3.6)$$

kde od roku 85 predpokladáme, že funkcia k_x je lineárna, teda

$$k_x = k_{85} - (x - k_{85})p, \quad (3.7)$$

kde k_{85} a p sú parametre. Odhad parametru p pritom dostaneme zo vzťahu

$$p = -\frac{[\log(m_{84}/m_{110}) + 26k_{85}]}{325}. \quad (3.8)$$

Model vyžaduje splnenie ďalších predpokladov. Prvým je nutnosť, aby miery úmrtnosti okolo veku 85 boli spoľahlivé a my sme tak mohli parameter k_{85} spočítať priamo z dát, čo v prípade Českej republiky môžeme považovať za splnené. Druhým predpokladom je stanovenie tzv. maximálne dosiahnuteľného veku, ktorý stanovili páni Coale a Kisker fixne na hodnotu 110 rokov. V praktickej časti sa pozrieme aj na iné hodnoty maximálne dosiahnuteľného veku, ale v tomto prípade budeme o výsledkoch daného modelu uvažovať s patričnou rezervou. Vek 110 bol stanovený na základe sledovania úmrtnosti v populácii Švédska a pravdepodobnosť úmrtnosti v tomto veku bola stanovená na hodnotu 1.0 pre mužov a 0.8 pre ženy z dôvodu, v tom čase ešte stále potrebného, explicitného rozdelenia úmrtnosti v rokoch najvyšších medzi pohlaviami. Výsledný model môžeme zapísať aj pomocou špecifickej miery úmrtnosti ako

$$m_x = \exp(ax^2 + bx + c), \quad (3.9)$$

kde a, b, c sú parametre modelu. Tento model sa v rámci Českej republiky nikdy oficiálne nepoužíval, v ostatných štátoch sveta, napr. vo Francúzsku, bol v minulosti uznávanou alteratívou pre Gompertz-Makehamov model, predovšetkým pre svoju numerickú stabilitu pri odhadovaní parametrov a pre všeobecne väčšiu flexibilitu modelu. V prípade zmeneného nastavenia maximálneho veku ω sa trajektórie vekovo špecifických mier úmrtnosti vo rokoch najvyšších vyrovnali v súlade s celkovou nádejou dožitia pri narodení jedinca [5].

3.3 Kannisto a Thatcherov model

Nasledujúce modely spadajú do kategórie logistických modelov, ktoré v poslednej dobe výrazne naberajú na popularite pri projekcii úmrtnosti do rokov najvyšších. Vo všeobecnosti možno uvažovať o logistickej funkcii v nasledujúcom tvare:

$$\mu_x = c + \frac{ae^{bx}}{1 + \alpha e^{bx}}. \quad (3.10)$$

Povšimnime si, že formulácia Makehamovho zákona je špeciálnym prípadom tejto rovnosti pre $\alpha = 0$. Malé hodnoty tohto parametru by tak vysvetľovali, prečo Gompertz-Makehamov model tak dobre fungoval. Vzhľadom na množstvo rozličných mien pre logistické modely budeme v tejto práci pracovať s tými, ktoré sú používané v rámci projektu Human Mortality Database, z ktorého čerpáme dáta pre praktickú analýzu. Thatcherov trojparametrický model má, v tom prípade, nasledujúci tvar

$$\mu_x = c + \frac{ae^{bx}}{1 + ae^{bx}}, \quad (3.11)$$

ktorý môžeme jednoduchou úpravou prepísať do tvaru

$$\text{logit}(q_x) = \log(a) + bx, \quad (3.12)$$

kde a , b , c sú parametre modelu a pre funkciu $\text{logit}(q_x)$ platí:

$$\text{logit}(q_x) = \log \frac{q_x}{p_x} = \log q_x - \log p_x. \quad (3.13)$$

Na podobnej myšlienke je založený aj Kannistov model, ktorý bol špeciálne navrhnutý pre vysoké roky a je v rámci spomínaného projektu používaný pre vyrovňovanie a extrapoláciu úmrtnostných tabuliek, keďže sa mu veľmi dobre darí popisovať súčasný trend vývoja úmrtnosti. Model má tvar

$$\mu_x = \frac{e^{a+b(x-80)}}{1 + e^{a+b(x-80)}}, \quad \text{pre } x > 80, \quad (3.14)$$

kde a , b sú nezáporné parametre modelu, pričom intenzita úmrtnosti vyrovnaná týmto modelom smeruje asymptoticky k 1. Model môžeme, podobne ako Thatcher, prepísať do nasledovného tvaru

$$\text{logit}(q_x) = ab(x - 80), \quad \text{pre } x > 80. \quad (3.15)$$

Oba modely sa teda líšia v zápise len minimálne, v rôznom pohľade na tvar lineárnej závislosti funkcie $\text{logit}(q_x)$ na veku a môžeme teda očakávať podobné výsledky, Kannistov model ale vo všeobecnosti predpovedá menšie pravdepodobnosti úmrtia v rokoch najvyšších. Obrovskou výhodou týchto modelov je ich flexibilita ohľadom maximálneho veku, modely dlhodobo vykazujú spoľahlivosť až do maximálneho veku $\omega = 120$. Výborné výsledky sú relevantné s momentálne prebiehajúcou diskusiou, či vôbec existuje fixná horná hranica dĺžky ľudského života. Pre roky 100 a viac sú to vo vyspelej Európe momentálne najvhodnejšie modely [17].

3.4 Helligman-Pollardov model

Najvšeobecnejší z predstavovaných modelov bol navrhnutý v roku 1980 a umožňuje modelovať úmrtnosť skrz všetky roky okrem toho nultého, kde každá aditívna časť sa zaoberá jednou, v úvode popísanou úmrtnosťou. To z neho robí jeden z najrozšírenejších a najviac používaných modelov v demografii. Vo svojej základnej podobe ide o 8 parametrický model v tvare:

$$q_x = A^{(x+b)^C} + De^{-E(\ln(x)-\ln(F))^2} + \frac{GH^x}{1 + GH^x}. \quad (3.16)$$

Pre účely nášho modelovania budeme pracovať s poslednou časťou, ktorá sa zaoberá problematikou najvyšších rokov, počet parametrov sa nám tak redukuje na 2. Parameter G pritom meria základný level úmrtnosti vo veku $x = 0$ a parameter H definuje mieru nárastu, ide o absolútny člen a sklon Gompertzovej krivky s oborom hodnôt $(0,1)$ a $(0, \infty)$ respektíve. Použijeme mierne upravený tvar:

$$q_x = \frac{be^{ax}}{1 + be^{ax}}, \quad (3.17)$$

kde a, b sú nové parametre modelu. Tento model je akýmsi kompromisom medzi vyššie uvedenými modelmi. Pravdepodobnosť úmrtia so zvyšujúcim sa vekom totižto narastá pomalšie ako u Gompertz-Makehama alebo Coale-Kiskera ale rýchlejšie ako u logistických modelov Kannista či Thatcher. Vhodnosť tohto modelu naprieč všetkými vekovými kategóriami bola doložená empirickými pozorovaniami úmrtnosti v Spojených štátoch amerických a Austrálii až do veku 90 rokov, opäť ako alternatíva ku Gompertz-Makehamovmu modelu [2].

3.5 Odhady parametrov

Pri odhadovaní parametrov pre jednotlivé modely budeme vychádzať z princípov nelineárnej regresie a predstavíme si podrobnejšie Gauss-Newtonov algoritmus, spolu s jeho praktickými modifikáciami, s ktorým budeme ďalej pracovať. V praktickej časti našej práce potom využijeme software DERAS (DEathRAtes-Simulation), kde sú nižšie popísané metódy implementované. Vychádzať budeme zo základnej publikácie [15].

3.5.1 Gauss-Newtonov algoritmus

Našou úlohou bude nájsť vektor parametrov $\boldsymbol{\theta}$ pomocou minimalizácie strednej kvadratickej odchýlky. Vektory a matice budeme ďalej značiť tučným písmom. Nech y_x sú skutočné, napozorované hodnoty úmrtí pre veky x a nech $f_x(\boldsymbol{\theta})$ sú vyrovnané hodnoty v príslušnom modeli s vektorovým parametrom $\boldsymbol{\theta}$. Potom je našou úlohou nájsť také $\boldsymbol{\theta}$, ktoré minimalizuje všeobecne funkciu

$$S(\boldsymbol{\theta}) = \sum_{x=1}^n [y_x - f_x(\boldsymbol{\theta})]^2 = \|\mathbf{y} - \mathbf{f}\|^2 = \sum_{x=1}^n r_x(\boldsymbol{\theta})^2, \quad (3.18)$$

kde symbolom $r_x(\boldsymbol{\theta})$ značíme rezíduá. Keďže sa v našej práci sústredíme na riziko dlhovekosti, budeme minimalizovať (3.18) pre hodnoty vekov od 60 do 85. Vrchná hranica je zvolená na základe dôveryhodnosti a dostatočného množstva dát pre Českú republiku v danom vysokom veku. Daný postup ale možno použiť pre ľubovoľné veky, preto pre jednoduchosť ponecháme v práci zvolené indexy.

Poznámka. Minimalizácia funkcie v (3.18) siaha aj za hranice nelineárnej regresie, v literatúrach o numerickej analýze sa často pracuje s výrazom $\|\mathbf{y} - \mathbf{f}\|^2$, v našej práci sa zameriavame najmä na reziduálne vyjadrenie $\sum_{x=1}^n r_x(\boldsymbol{\theta})$.

Vychádzať budeme z lineárnej aproximácie funkcie $f_x(\boldsymbol{\theta})$ Taylorovým rozvojom v bode $\boldsymbol{\theta}^{(a)}$, teda

$$f_x(\boldsymbol{\theta}) \approx f_x(\boldsymbol{\theta}^{(a)}) + \frac{\partial f_x(\boldsymbol{\theta}^{(a)})}{\partial \boldsymbol{\theta}^T} (\boldsymbol{\theta} - \boldsymbol{\theta}^{(a)}), \quad (3.19)$$

kde symbolom T budeme značiť transpozíciu vektoru alebo matice. Dosadením aproximácie 3.19 do 3.18 dostávame úlohu minimalizovať výraz

$$\|\mathbf{r}^{(a)} - \mathbf{F}^{(a)}(\boldsymbol{\theta} - \boldsymbol{\theta}^{(a)})\|^2, \quad (3.20)$$

kde $\mathbf{r}^{(a)} = \mathbf{y} - \mathbf{f}(\boldsymbol{\theta}^{(a)})$ a $\mathbf{F}^{(a)}$ je $\partial \mathbf{f} / \partial \boldsymbol{\theta}^T$ vyčíslená v bode $\boldsymbol{\theta}^{(a)}$. Úloha 3.20 má riešenie

$$\boldsymbol{\theta} - \boldsymbol{\theta}^{(a)} = (\mathbf{F}^{(a)T} \mathbf{F}^{(a)})^{-1} \mathbf{F}^{(a)T} \mathbf{r}^{(a)}, \quad (3.21)$$

ktoré vedie ku Gauss-Newtonovmu algoritmu

$$\boldsymbol{\theta}^{(a+1)} = \boldsymbol{\theta}^{(a)} + \boldsymbol{\delta}^{(a)}, \quad (3.22)$$

kde

$$\boldsymbol{\delta}^{(a)} = (\mathbf{F}^{(a)T} \mathbf{F}^{(a)})^{-1} \mathbf{F}^{(a)T} \mathbf{r}^{(a)}. \quad (3.23)$$

Aby sme pochopili hlavné prednosti a slabiny uvedeného postupu, pozrime sa na vzťah Gauss-Newtonovho algoritmu a Newtonovho algoritmu. Gradient a Hessián funkcie $S(\boldsymbol{\theta})$ majú tvar

$$g(\boldsymbol{\theta}) = \frac{\partial S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 2 \sum_{x=1}^n r_x(\boldsymbol{\theta}) \frac{\partial r_x(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 2 \mathbf{J}^T \mathbf{r}, \quad (3.24)$$

respektíve

$$\mathbf{H}(\boldsymbol{\theta}) = \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = 2 \sum_{x=1}^n \frac{\partial r_x(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial r_x(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} + 2 \sum_{x=1}^n r_x(\boldsymbol{\theta}) \frac{\partial^2 r_x(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = 2(\mathbf{J}^T \mathbf{J} + \mathbf{A}), \quad (3.25)$$

kde

$$\mathbf{J} = \mathbf{J}(\boldsymbol{\theta}) = \left[\left(\frac{\partial r_x}{\partial \theta_j} \right) \right] = \frac{\partial \mathbf{r}}{\partial \boldsymbol{\theta}^T} = -\mathbf{F}(\boldsymbol{\theta}) \quad (3.26)$$

a

$$= \mathbf{A}(\boldsymbol{\theta}) = \sum_{x=1}^n r_x(\boldsymbol{\theta}) \frac{\partial^2 r_x(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}. \quad (3.27)$$

Jeden krok Newtonovho algoritmu je

$$\boldsymbol{\delta}^{(a)} = -\mathbf{H}^{(a)-1} \mathbf{g}^{(a)} = -(\mathbf{J}^{(a)T} \mathbf{J}^{(a)} + \mathbf{A}^{(a)})^{-1} \mathbf{J}^{(a)T} \mathbf{r}^{(a)}. \quad (3.28)$$

Porovnaním s (3.23) ($\mathbf{J}^{(a)} = -\mathbf{F}^{(a)}$) vidíme, že Gauss-Newtonov algoritmus vznikne miernou modifikáciou Newtonovho algoritmu vynechaním časti Hessiánu, konkrétne $\mathbf{A}(\boldsymbol{\theta})$ z (3.27).

Hlavnou výhodou tohto postupu je fakt, že vyžaduje len prvé derivácie $r_x(\boldsymbol{\theta})$ uložené v \mathbf{J} , nie je potrebné ukladať žiadnu inú ďalšiu aproximáciu Hessiánu. V prípade, že člen $\hat{\mathbf{A}} = \mathbf{A}(\hat{\boldsymbol{\theta}})$ je malý v porovnaní s členom $\hat{\mathbf{J}}^T \hat{\mathbf{J}}$ vo výraze (3.25), počty krokov u oboch algoritmov sú približne rovnaké, pričom ale každý jeden krok Gauss-Newtonovho iteračného algoritmu je časovo a aj pamäťovo menej náročný než u Newtonovho algoritmu. Pre naše dáta je táto požiadavka splnená vzhľadom na dôveryhodnosť dát, ktoré nevykazujú väčšie výkyvy v rokoch 60 až 85. Reziduá v modeloch sú malé, čo zodpovedá požiadavke, aby bolo $\hat{\mathbf{A}}$ zanedbateľné v porovnaní s $\hat{\mathbf{J}}^T \hat{\mathbf{J}}$ podľa [4].

Nevýhodou spomínaného postupu je fakt, že časť Hessiánu $\mathbf{A}(\boldsymbol{\theta})$ môže byť významná, v tomto prípade je nutné použiť postupy pre riešenia problémov s

tzv. veľkými rezíduami. Ako príklad môžeme uviesť postup založený na kvázi-Newtonovej aproximácii $\mathbf{A}(\boldsymbol{\theta})$ alebo Gill-Murrayho metódu, pre podrobnosti pozri publikáciu [4]. Problemom môže byť aj situácia keď matica $\mathbf{J}^{(a)}$ nemá plnú hodnotu, či jej zlá podmienenosť, čo môže viesť k lokálnej divergencii algoritmu. Preto sa v nasledujúcej sekcii pozrieme na praktické modifikácie tohto algoritmu, ktoré povolujú singularitu alebo zlú podmienenosť matíc $\mathbf{J}^{(a)T}\mathbf{J}^{(a)}$.

3.5.2 Levenberg-Marquadtové metódy

Vydeme z modifikácie kroku Gauss-Newtonovej metódy (3.24) voľbou

$$\boldsymbol{\delta}^{(a)} = -(\mathbf{J}^{(a)T}\mathbf{J}^{(a)} + \eta^{(a)}\mathbf{D}^{(a)})^{-1}\mathbf{J}^{(a)T}\mathbf{r}^{(a)}, \quad (3.29)$$

kde $\mathbf{D}^{(a)}$ je diagonálna matica s kladnými prvkami na diagonále s tým, že matica $\mathbf{J}^T\mathbf{J} + \eta\mathbf{D}$ je lepšie podmienená než matica $\mathbf{J}^T\mathbf{J}$. Populárnymi voľbami matice $\mathbf{D}^{(a)}$ sú matica diagonálna, to jest $\mathbf{D}^{(a)} = \mathbf{I}_p$ (p je všeobecne hodnota matice modelu), alebo matica, ktorej prvky na diagonále sú totožné s prvkami na diagonále $\mathbf{J}^{(a)T}\mathbf{J}^{(a)}$, čím zostane metóda aproximatívne invariantná v prípade zmeny škály $\boldsymbol{\theta}$. Pre úplnosť dodajme, že krok $\boldsymbol{\delta}^{(a)}$ nie je vo väčšine dobrých softwarov spočítaný na základe výpočtu $\mathbf{J}^{(a)T}\mathbf{J}^{(a)}$, ale získa sa ako riešenie lineárneho problému minimalizáciou $\boldsymbol{\delta}$ v podobe metódy najmenších štvorcov

$$\underset{\boldsymbol{\delta}}{\text{Min}} \left\| \begin{pmatrix} \mathbf{r}^{(a)} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{J}^{(a)} \\ (\eta^{(a)}\mathbf{D}^{(a)})^{1/2} \end{pmatrix} \boldsymbol{\delta} \right\|^2 \quad (3.30)$$

pomocou numericky stabilných metód.

V prípade, že zvolíme $\mathbf{D}^{(a)} = \mathbf{I}_p$, smer $\boldsymbol{\delta}^{(a)}$ vo vzťahu (3.29) interpoluje medzi Gauss-Newtonovým algoritmom ($\eta^{(a)} \rightarrow 0$) a najstrmším možným poklesom ($\eta^{(a)} \rightarrow \infty$), kedy krok dĺžky $\|\boldsymbol{\delta}^{(a)}\|$ konverguje k nule. Pre $\eta^{(a)} > 0$ je naša matica $\mathbf{J}^{(a)T}\mathbf{J}^{(a)} + \eta^{(a)}\mathbf{D}^{(a)}$ pozitívne definitná, keďže $\mathbf{D}^{(a)}$ je pozitívne definitná matica. Dá sa ukázať, že pre $\eta^{(a)} \rightarrow \infty$ je $\boldsymbol{\delta}^{(a)}$ skutočne poklesom pri hľadaní minimalizačnej funkcie, viď publikáciu [15], sekciu 13.2.3b pre podrobnosti. Vzhľadom na to, že pre $\eta^{(a)} \rightarrow \infty$ konverguje $\boldsymbol{\delta}^{(a)}$ k nule, môžeme voľbou dostatočne veľkého $\eta^{(a)}$ redukovať rezíduu v (3.18). Príliš veľkou voľbou $\eta^{(a)}$ ale riskujeme, že algoritmus bude potrebovať veľa krokov a bude tak veľmi pomalý. Levenberg-Marquadtové metódy sa líšia práve v rôznom prístupe k voľbe a aktualizácii $\eta^{(a)}$.

Počiatočným pokusom Levenberga bola voľba $\eta^{(a)}$ získaná minimalizáciou $S(\boldsymbol{\theta})$ pre $\boldsymbol{\theta}$ v tvare $\boldsymbol{\theta}^{(a)} + \boldsymbol{\delta}^{(a)}$, kde $\boldsymbol{\delta}^{(a)}$ je dané (3.29). Tento postup ale nefungoval, keďže každý pokus o získanie hodnoty $\eta^{(a)}$ vyžadoval náročné riešenie ďalšieho problému tvaru (3.30). S riešením prišiel Marquadt, ktorý navrhol nasledujúci postup. Začneme na malej pozitívnej hodnote, napr. $\eta^{(1)} = 0.01$. V prípade, že krok $\boldsymbol{\delta}^{(a)}$ tvaru (3.29) redukuje v a -tej iterácii $S(\boldsymbol{\theta})$, položíme $\boldsymbol{\theta}^{(a+1)} = \boldsymbol{\theta}^{(a)} + \boldsymbol{\delta}^{(a)}$ a vydáme $\eta^{(a)}$ určitým faktorom, napr. $\eta^{(a+1)} = \eta^{(a)}/10$. Tým dosiahneme priblíženie algoritmu ku Gauss-Newtonovmu algoritmu, a zároveň posunieme algoritmus ku väčšiemu kroku v ďalšej iterácii (3.29). V prípade, že sa nám v a -tej iterácii nepodarí pomocou $\boldsymbol{\delta}^{(a)}$ redukovať $S(\boldsymbol{\theta})$, zvýšime hodnotu $\eta^{(a)}$, napr. $\eta^{(a)} \rightarrow 10\eta^{(a)}$ a prepočítame $\boldsymbol{\delta}^{(a)}$. Tento postup opakujeme, až kým nedosiahneme redukciiu $S(\boldsymbol{\theta})$.

Minimalizačný problém (3.30) musíme teda riešiť len v niektorých iteráciach. Tento postup použijeme aj v našej práci.

V súčasnej dobe existujú nové modifikácie Levenberg-Marquadtovho algoritmu založené na rôznych obmedzeniach jednotlivých krokov, tzv. *trust-region* metódy, ktoré pomáhajú riešiť popísané nedostatky, pre podrobnosti pozri [15], sekcia 13.3.2. Pre naše dáta je ale algoritmus postačujúci, pre celkový počet 1 440 vyrovnaných a extrapolovaných tabuliek zlyhal algoritmus len raz a aj to pri vysoko volatilných dátach pre mužov v 50. rokoch pri vekovej hranici 115 rokov.

3.6 Intervaly spoľahlivosti

V praktickej časti tejto práce budeme pracovať s tromi intervalmi spoľahlivosti. Intervaly spoľahlivosti pre odhadnuté parametre môžeme získať rovnako ako u lineárnej regresie z predpokladu normálneho, resp. studentovho t -rozdelenia. Pre konštrukciu intervalov spoľahlivosti pre regresnú krivku až do veku 80 môžeme rovnako použiť štandardné intervaly spoľahlivosti z lineárnej regresie. Tieto roky boli použité pri odhadovaní parametrov, máme pre ne dostatok dát a predpoklad normálneho rozdelenia rezíduí je tu zrejme splnený. Cieľom tejto práce nie je zhrnúť základy lineárnej regresie (pozri [14]), preto ich v práci nebudeme uvádzať a zameriame sa na problematiku intervalov spoľahlivosti pre roky najvyššie.

Klasický postup z lineárnej regresie nemôžeme využiť pre veku nad 85 rokov, ktoré vykazujú značný nárast rozptylu s rastúcim vekom a v prípade projekcie do rokov najvyšších by bol predpoklad o normalite rezíduí hrubo porušený. Informácie o úmrtiach a o počte ľudí navyiac nemajú vplyv na výsledný model. Modelovanie úmrtnosti pre roky najvyššie je predpoveďou modelov pre ďalšie časové (vekové) okamihy. Využijeme preto špeciálny prístup pre všeobecne rastové (v našom prípade pod pojmom rast chápeme nárast veku jedinca) krivky popísaný napr. v [10].

3.6.1 Predikčné intervaly pre rastovú krivku

Nadalej budeme pracovať s 3 parametrovou rastovou krivkou v tvare:

$$r_x = f_x(a, b, c) + \epsilon_x, \quad (3.31)$$

kde a , b , c sú parametre krivky, x je skúmaný vek a ϵ_x je aditívna chybová zložka. Môžeme sa presvedčiť, že všetky skúmané krivky sú skutočne tohto tvaru (v niektorých sa len vynecháva nadbytočný parameter c , v súlade s minulým zápisom $\theta = (a, b, c)$).

Predpovedný interval do času $x+\tau$ založíme na podmienenej strednej hodnote, $E(r_{x+\tau}|r_x, \dots, r_0)$ a podmienenom rozptyle, $\text{var}(r_{x+\tau}|r_x, \dots, r_0)$. Podmienenú strednú hodnotu získame extrapoláciou z modelu, ale podmienený rozptyl je treba najskôr spočítať. Za predpokladu nezávislosti medzi hodnotami rastovej krivky a chybovou zložkou je možné napísať podmienený rozptyl v tvare:

$$\text{var}(r_{x+\tau}|y_x, \dots, y_0) = \text{var}(f_{x+\tau}(a, b, c)|r_x, \dots, r_0) + \text{var}(\epsilon_{x+\tau}|r_x, \dots, r_0). \quad (3.32)$$

Pre úplnosť dodajme, že v našom prípade budeme pod časovým indexom 0 rozumieť vek 60, pod časovým indexom x vek 85 a predpovede budeme uvažovať v niektorých prípadoch až do veku 120, to jest $\tau = 1, 2, \dots, 35$. Označme chyby pri odhadoch parametrov a, b, c postupne $\epsilon_a, \epsilon_b, \epsilon_c$ a predpokladajme, že vektor $(\epsilon_a, \epsilon_b, \epsilon_c)$ má nulovú strednú hodnotu a kovariančná matica vzniká ako vedľajší produkt pri odhadovaní parametrov, to jest

$$\Sigma = \text{var} \begin{bmatrix} \epsilon_a \\ \epsilon_b \\ \epsilon_c \end{bmatrix} = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} & \sigma_{ac} \\ \sigma_{ab} & \sigma_b^2 & \sigma_{bc} \\ \sigma_{ac} & \sigma_{bc} & \sigma_c^2 \end{bmatrix}. \quad (3.33)$$

Skutočná hodnota odhadnutého parametru a , použitá pri výpočte predikcie v čase x , je $a + \epsilon_a$, index τ vynecháme pre jednoduchosť. Podobne pre parametre b a c . Podmienový rozptyl chybovej zložky ϵ_x označíme σ_ϵ^2 a rovnako ako pri kovariančnej matici (3.33) ho získame ako vedľajší produkt pri odhadoch parametrov. Postup pri odhadovaní parametrov bol popísaný v sekcii 2.5., pričom kovariančnú maticu medzi odhadmi θ_i a θ_j získame z

$$\frac{2S(\boldsymbol{\theta})}{N - p} \mathbf{H}(\boldsymbol{\theta}^{-1})_{ij}. \quad (3.34)$$

K odhadu rozptylu rastovej funkcie využijeme prístup založený na jeho aproximácii pomocou Taylorovho rozvoja:

$$\begin{aligned} \text{var}(f_{x+\tau}(a, b, c) | r_x, \dots, r_0) &= \left(\frac{\partial \mathbf{f}}{\partial a} \right)^2 \sigma_a^2 + \left(\frac{\partial \mathbf{f}}{\partial b} \right)^2 \sigma_b^2 + \left(\frac{\partial \mathbf{f}}{\partial c} \right)^2 \sigma_c^2 \\ &+ 2 \frac{\partial \mathbf{f}}{\partial a} \frac{\partial \mathbf{f}}{\partial b} \sigma_{ab} + 2 \frac{\partial \mathbf{f}}{\partial a} \frac{\partial \mathbf{f}}{\partial c} \sigma_{ac} + 2 \frac{\partial \mathbf{f}}{\partial b} \frac{\partial \mathbf{f}}{\partial c} \sigma_{bc}, \end{aligned} \quad (3.35)$$

pre podrobnosti pozri [16]. S využitím (3.35) a (3.32) dostávame $100(1 - \alpha)\%$ predikčný interval pre $r_{x+\tau}$ v tvare:

$$E(r_{x+\tau} | r_x, \dots, r_0) \pm k_{\alpha/2} (\text{var}(r_{x+\tau} | r_x, \dots, r_0))^{1/2}, \quad (3.36)$$

kde $k_{\alpha/2}$ je vhodná hodnota náhodnej premennej vysvetľujúca chybu predpovede. Najčastejšie sa predpokladá, že táto náhodná premenná je normálne rozdelená. Alternatívnym predpokladom môže byť studentovo t rozdelenie alebo hodnota $\alpha^{-1/2}$ plynúca z Chebyshevovej nerovnosti, viď [10] pre podrobnosti. V nasledujúcej kapitole si ukážeme popísaný postup konštrukcie predikčného intervalu pre konkrétny model.

Poznámka. Okrem Taylorovej aproximácie sú v literatúre [10] spomínané aj iné možné prístupy, založené na explicitnej hustote chýb odhadov ϵ_a, ϵ_b a ϵ_c , či predikčné intervaly založené na *bootstrappingu*. Prvá spomínaná metóda je vhodná pri dátach, kde máme k dispozícii len málo pozorovaní. Metóda bootstrappingu sa spomedzi týchto metód ukázala ako príliš striktná, a teda najmenej vhodná pre rastové funkcie, preto budeme využívať prístup založený na Taylorovej aproximácii rozptylu.

Kapitola 4

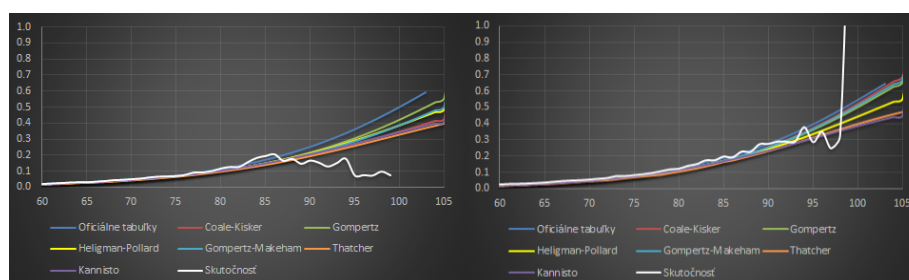
Analýza úmrtnosti

V nasledujúcej kapitole aplikujeme predstavené modely na populáciu Českej republiky v priebehu posledných 65 rokov, interpretované hlavné výsledky budú z priloženého pracovného súboru *úmrtnostné tabuľky.xlsm*. V prípade, že by čitateľa zaujímali podrobnosti ohľadom interpretovaných čísel, detailov, grafov, tabuľiek alebo snaha o hlbšie pochopenie, je nahliadnutie do tohto súboru kľúčové. Na prvom liste *Guide* je popísaná štruktúra a význam jednotlivých listov. Pri grafickej prezentácii výsledkov budeme pod pojmom skutočnosť uvažovať aproximáciu pravdepodobností úmrtia podľa vzťahu (2.5).

Na úvod sa pozrieme na vývoj úmrtnosti v rokoch 1950 až 1999 a preskúmame vývoj jednotlivých parametrov v čase. Následne sa budeme zaoberať situáciou v novom tisícročí, kde si ukážeme trendy vývoja úmrtnosti vo vysokých rokoch 80 a 90, posúdime vhodnosť jednotlivých modelov na základe vyrovnania pravdepodobností úmrtia v rokoch 60 až 85, a následne aj ich projekcie do rokov najvyšších. V závere si ukážeme postup pre konštrukciu intervalov spoľahlivosti pre projektované veki na príklade z dát.

4.1 Úmrtnosť v rokoch 1950 až 1999

U mužov je situácia v tomto období veľmi komplikovaná. Od roku 1950 až do roku 1975 boli u mužov pravdepodobnosti úmrtia pre vysoké veki z roka na rok značne odlišné a vysoko volatilné, najmä pre stále relatívne nízky počet osôb dožívajúcich sa vysokých rokov. Obrázok 3.1 zachytáva skutočné pravdepodobnosti úmrtia v dvoch po sebe idúcich rokoch 1960 a 1961, spolu s oficiálnymi úmrtnostnými tabuľkami a odhadnutými modelmi.

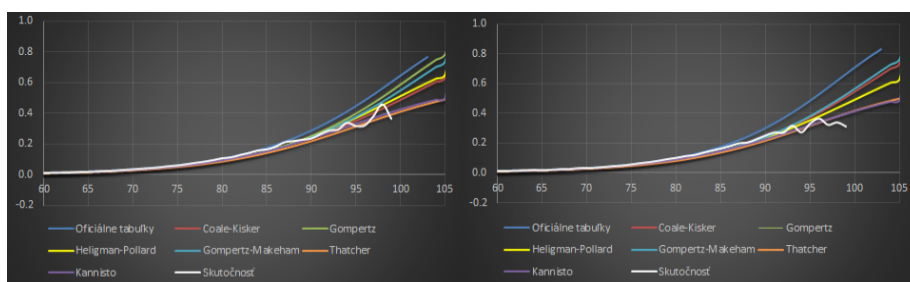


Obr. 4.1: Pravdepodobnosti úmrtia v rokoch 1960 a 1961, muži

Pravdepodobnosť úmrtia vo veku 80 rokov sa v tom čase pohybovala okolo 13%, dnes je táto pravdepodobnosť na hodnote okolo 7%. Pri pohľade na počet žijúcich ľudí v rokoch 1950 a 1970 máme 8 082, respektíve 8 928 ľudí vo veku 80 rokov a len 599, respektíve 980 ľudí žijúcich vo veku 90 rokov. V porovnaní so súčasnosťou, kde je to 19 734 ľudí vo veku 80 a 3 819 ľudí vo veku 90, je jasné, o koľko menej pozorovaní sme v tom čase mali k dispozícii. Až v roku 1998 bolo v Českej republike viac ako 2 000 mužov vo veku 90 rokov, u žien v roku 1968.

Od roku 1975 už začínajú dáta poskytovať prvé spoľahlivé výsledky, Gompertz-Makehamov model spolu s oficiálnymi úmrtnostnými tabuľkami Českého štatistického úradu v rokoch najvyšších často nadhodnocujú skutočné pravdepodobnosti úmrtia s výnimkou rokov 1985 až 1990, kedy bola úmrtnosť mužov v pokročilom veku relatívne vysoká. Vo väčšine rokov sú z tejto doposiaľ grafickej analýzy najvhodnejšie logistické modely, do úvahy je ale treba brať vysokú volatilitu dát.

U žien je veľmi zaujímavý výrazný nárast počtu žien vo veku 90 rokov. V období od roku 1950 do roku 1975 sa tento počet takmer strojnásobil (z 885 na 2 548), v roku 1968 máme k dispozícii údaje o 17 581 ženách vo veku 80 rokov a o 2 099 ženách vo veku 90 rokov, najmä vďaka zníženej úmrtnosti žien v 60. rokoch 20. storočia. Do tohto obdobia, to jest do roku 1960, bol Gompertz-Makehamov model pre ženy skutočne najvýhodnejší, z grafickej prezentácie modelov v priloženom súbore je to zrejmé. Od roku 1960 môžeme ale pozorovať postupné znižovanie pravdepodobností úmrtia, kedy sú pravdepodobnosti úmrtia týmto modelom nadhodnocované v prakticky každom skúmanom roku a do popredia sa postupom času dostávajú logistické modely a model Heligman-Pollarda. Obrázok 3.2 zachytáva situáciu v rokoch 1974 a 1984, exponenciálne modely, vrátane toho Gompertz-Makehamovho začínajú byť pomaly nevyhovujúce.

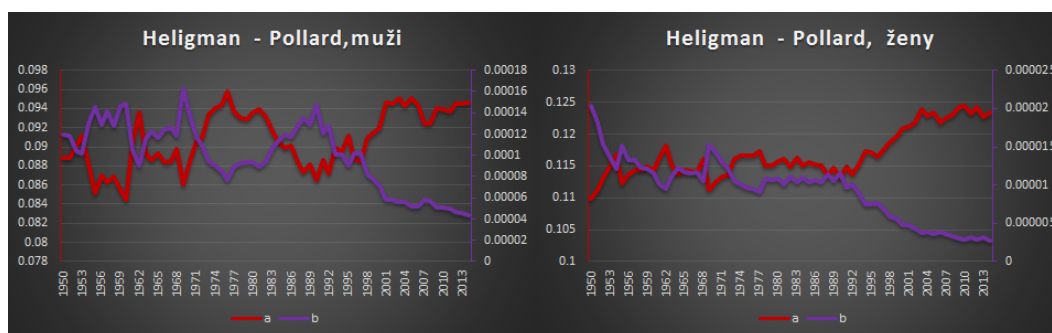


Obr. 4.2: Pravdepodobnosti úmrtia v rokoch 1974 a 1984, ženy

Zaujímavá je aj určitá paralela vo vývoji úmrtnosti medzi mužmi a ženami, úmrtnosť žien má miestami podobné charakteristiky ako úmrtnosť mužov niekoľko rokov dozadu, čo zodpovedá presvedčeniu, že samotný vývoj úmrtnosti nie je špecifikovaný pohlavím osoby.

Poznámka. Ku koncu 90. rokov 20. storočia sú špecifické miery úmrtnosti u žien v rokoch najvyšších vo svojich historických minimách. V tomto krátkom časovom rozpätí 6 rokov sa totižto počet obyvateľov vo vysokých rokoch výrazne zvýšil, čo môže byť vysvetlené silnou dožívajúcou generáciou alebo výrazným zlepšením starostlivosti o seniorov v zdravotníctve.

Významnú odlišnosť volatility dát medzi jednotlivými pohlaviami a rozličné trendy klesajúcich úmrtností si môžeme všimnúť aj pri pohľade na časový vývoj odhadnutých parametrov. Obrázok 3.3 zachytáva vývoj odhadov parametrov pre Heligman-Pollardov model. Divoký priebeh u mužov je spojený práve s volatilitou ich dát, od roku 1998 u mužov (resp. od roku 1992 u žien) pozorujeme ďaleko hladší priebeh. Pravdepodobnosti úmrtia pre zvolený vek 85 rokov u (3.17) môžeme takisto spočítať: u mužov pozorujeme v rámci obdobia 1950 až 1999 pokles z 18.5% na 15% a u žien z 18.5% na 12.5%, čo predstavuje značný rozdiel.

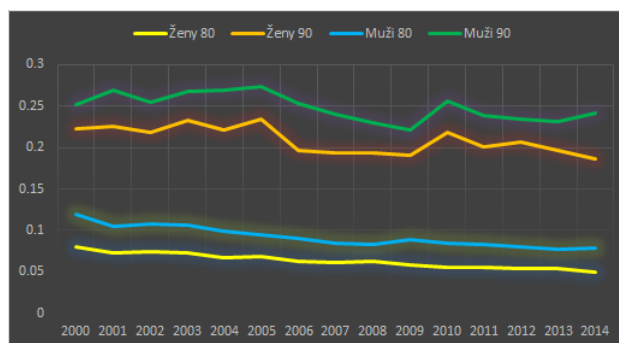


Obr. 4.3: Odhady parametrov v jednotlivých kalendárnych rokoch

4.2 Vývoj úmrtnosti v posledných 15 rokoch

V dôsledku neustále sa rozvíjajúceho technického pokroku v medicíne, sociológii, farmácii a iných, so smrťou priamo, či nepriamo súvisiacich vied, sa kvalita života zlepšuje každým rokom, čo sa samozrejme odráža aj na vývoji úmrtnosti, v poslednej dobe je tento pokrok veľmi výrazný. Ak sa napríklad pozrieme na počet žijúcich osôb vo veku 90, u žien pozorujeme v roku 2014 58 % nárast oproti roku 2000 (z 6 824 na 10 748), u mužov až 60 % percentný (z 2 388 na 3 819). Vynára sa tak prirodzene otázka, či Gompertz-Makehamova krivka nenahodnocuje pravdepodobnosti umierania, ako je tomu vo vyspelejších krajinách sveta.

Pre demografiu je dôležitým ukazovateľom špecifická miera úmrtnosti, ktorá udáva presnejší pohľad na vývoj úmrtnosti než počty žijúcich. Obrázok 3.4 zachytáva tento ukazovateľ za posledných 15 rokov.



Obr. 4.4: Vekovo špecifická miera úmrtnosti

Môžeme si všimnúť, že špecifická miera úmrtnosti pre vek 80 zrejme klesá, u žien z 0.081 na 0.050 a u mužov z 0.119 na 0.078. Vo veku 90 pozorujeme ďaleko väčšie výkyvy a menej hladký priebeh, u žien je ale klesajúci trend dobre viditeľný, u mužov sú tieto výkyvy spojené s menším počtom dát a ich väčšou volatilitou, ktorú môžeme pozorovať aj v iných členských štátoch EÚ, mierne klesajúci trend je aj v tomto prípade zrejímavý.

Ako sme sa mohli presvedčiť, špecifická miera úmrtnosti pre veky 80 a 90 pre Českú republiku v poslednom období klesá, podobne ako pre ostatné vysoké roky, čo zodpovedá klesaniu pravdepodobnosti úmrtia podľa (2.2). Gompertz-Makehamov model tak môže skutočne nadhodnocovať pravdepodobnosti úmrtia pre vysoké roky, má teda zmysel pozrieť sa na ostatné modely, u ktorých je konvergencia pravdepodobnosti umierania v rokoch najvyšších pomalšia. Vhodnými kandidátmi sú práve modely Heligman-Pollarda, či logistické modely.

4.3 Vhodnosť jednotlivých modelov na populáciu Českej republiky

V tejto časti preskúmame správanie sa jednotlivých modelov v populácii Českej republiky. Porovnáme kvalitu vyrovnaní modelov pre roky 60 až 85, preskúmame projekčné vlastnosti modelov pre roky najvyššie a pozrieme sa aj na voľbu maximálneho veku. Pre posúdenie kvality vyrovnaní definujeme strednú tvorcovú chybu pre vyrovnané pravdepodobnosti úmrtia \hat{q}_x ako

$$MSE = \sum_{x=60}^{85} \frac{\left(\left(m_x - \frac{m_x^2}{2} + \frac{m_x^3}{6} \right) - \hat{q}_x \right)^2}{26}. \quad (4.1)$$

Tento ukazovateľ nadobúda v priebehu posledných 15 rokov u mužov malé hodnoty, ktoré sa v rámci jednotlivých modelov líšia len minimálne. Thatcherov model spolu so základným Gompertz-Makehamovým a Coale-Kiskerovým modelom vyrovnáva dáta o trochu presnejšie. To môžeme demonštrovať napr. rokom 2008, kedy sme napočítali hodnoty

Model	MSE
Coale-Kisker	4.56 10^{-6}
Gompertz	10.72 10^{-6}
Heligman-Pollard	13.10 10^{-6}
Kannisto	14.65 10^{-6}
Gompertz-Makeham	4.03 10^{-6}
Thatcher	3.77 10^{-6}

Obr. 4.5: Tabuľka 1: MSE, muži, 2008

U žien pozorujeme väčšiu variabilitu modelov pri vyrovnaní dát, ale aj v rámci samotných modelov, avšak situácia sa v podstate mení v podstate z roka na rok a je ťažké vyvodzovať závery ohľadom porovnávania modelov. Modely sú

si navzájom dobre konkurujúce a o žiadnom nemôžeme prehlásiť, že by nebol vhodný. Tabuľka 2 zachytáva situáciu z roku 2008 pre porovnanie s mužmi. V priloženom súbore možno nájsť vývoj MSE v priebehu posledných 15 rokov.

Model	MSE
Coale-Kisker	$8.57 \cdot 10^{-5}$
Gompertz	$2.94 \cdot 10^{-5}$
Heligman-Pollard	$2.53 \cdot 10^{-5}$
Kannisto	$2.30 \cdot 10^{-5}$
Gompertz-Makeham	$7.06 \cdot 10^{-5}$
Thatcher	$3.81 \cdot 10^{-5}$

Obr. 4.6: Tabuľka 2: MSE, ženy, 2008

Dôležitým identifikačným nástrojom pre vhodnosť vybraného modelu je správanie sa reziduálnej zložky, ktorú získame ako

$$\epsilon_x = \hat{q}_x - \left(m_x - \frac{m_x^2}{2} + \frac{m_x^3}{6} \right). \quad (4.2)$$

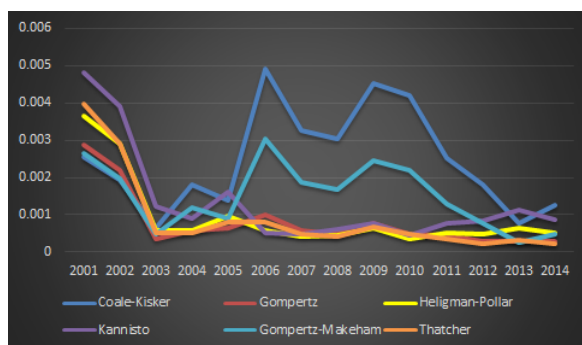
Tá nám pomôže odhaliť prípadné nadhodnocovanie, či podhodnocovanie reálnych pravdepodobností úmrtia, najmä u vyšších rokov. U rokov 60 až 85 sme tieto trendy odpozorovali len veľmi minimálne, pri Coale-Kiskerovom a Gompertzov-Makehamovom modeli boli tieto pravdepodobnosti jemne nadhodnotené, čo vysvetľuje vyššie hodnoty strednej štvorcovej chyby pri vyrovnávaní. Stredná hodnota v (4.2) bola u ostatných modelov skoro vždy nulová, až do rádu 10^{-4} . Získané výsledky sú aj pri pohľade na reziduálnu zložku uspokojivé, navzájom veľmi podobné a o žiadnom modeli sa nedá prehlásiť, že by bol pre účely vyrovnania daných rokov vyslovene nevhodný.

Poznámka. V druhej kapitole sme sa v rámci Gompertz-Makehamovho modelu venovali aj jeho pôvodnej konštrukcii pre roky 20 až 60 rokov. Teraz sme ale ukázali, že tento model sa javí ako dobrý kandidát aj pre vyrovnanie rokov 60 až 85. Tento posun pôvodných rokov na vyššiu hranicu jasne ukazuje smer vývoja, ktorým úmrtnosť za posledných 100 až 150 rokov prešla.

Pozrime sa teraz na roky najvyššie, kde jasne uvidíme rozličné správanie sa vybraných modelov. Pre účely projekčných vlastností jednotlivých modelov definujeme strednú štvorcovú chybu projekcie

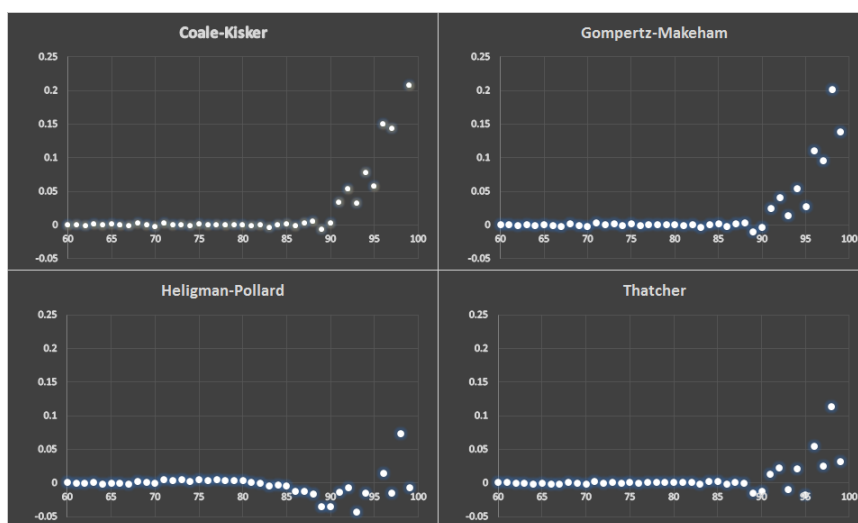
$$MSEP = \sum_{x=60}^{99} \frac{\left(\left(m_x - \frac{m_x^2}{2} + \frac{m_x^3}{6} \right) - \hat{q}_x \right)^2}{40}. \quad (4.3)$$

Tento ukazovateľ nám poskytuje prvý náhľad na správanie skúmaných modelov v rámci celého spektra skúmaných vekov. Vek 99 je maximálny vek, pre ktorý máme v dátach údaje o počte žijúcich k 1. 7. daného roku a o počte mŕtvych, teda údaje o špecifickej miere úmrtnosti. Najkôr sa pozrieme na mužov. Obrázok 3.5 zachytáva vývoj MSEP za posledných 14 rokov. Rok 2000 neuvádzame z interpretačného dôvodu, napozorovali sme vysoké hodnoty tohto kritéria.



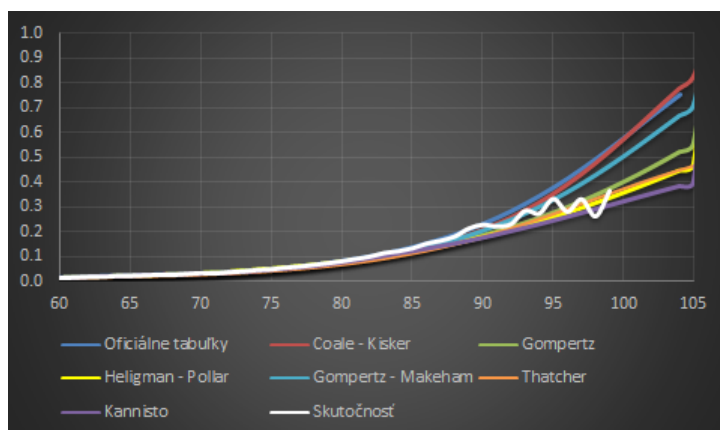
Obr. 4.7: Vývoj strednej štvorcovej chyby projekcie, muži

Je zrejmé, že modely Coale-Kiskera a Gompertz-Makehama podávajú najmenej uspokojivé výsledky, naopak Heligman-Pollardov a Thatcherov model sa zdajú byť v tomto období u mužov najvhodnejšími. Pri analýze reziduálnej zložky zistíme, že Coale-Kiskerov a Gompertz-Makehamov model značne nadhodnocujú pravdepodobnosti úmrtia v rokoch najvyšších, na obrázku 3.6 je zachytená napríklad situácia z roku 2010. Podobný vývoj sme pozorovali aj v ostatných skúmaných rokoch, stredná hodnota (4.2) pre veku 60 až 99 sa u modelov Coale-Kiskera a Gompertz-Makehama často pohybuje okolo hodnoty 0.02, u ostatných modeloch je vo väčšine prípadov v absolútnej hodnote menšia ako 0.01.



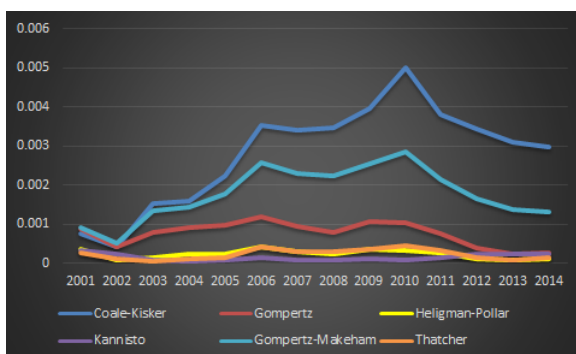
Obr. 4.8: Reziduá, muži, 2010

Na základe kritéria (4.3) a analýzy reziduálnej zložky sme zistili, že modely Gompertz-Makehama a Coale-Kiskera často nadhodnocujú pravdepodobnosti úmrtia u mužov v rokoch najvyšších. Podobný trend sme zaznamenali aj v druhej polovici 20. storočia, tu sme ale vzhľadom na nedostatok dát nepodrobili modely väčšej kritike. V posledných rokoch ale logistické modely, či model Heligman-Pollarda, vykazujú ďaleko lepšie výsledky pre roky najvyššie, na obrázku 3.9 je zachytená napríklad situácia z roku 2009. Zaujímavé ale je, že okolo veku 90 tieto modely miestami mierne podhodnocujú pravdepodobnosti úmrtia, naopak Gompertz-Makehamov a Coale-Kiskerov model ich zachytia spoľahlivo. Do veku 90 sú tieto modely ešte stále uspokojivé.



Obr. 4.9: Pravdepodobnosti úmrtia, muži, 2009

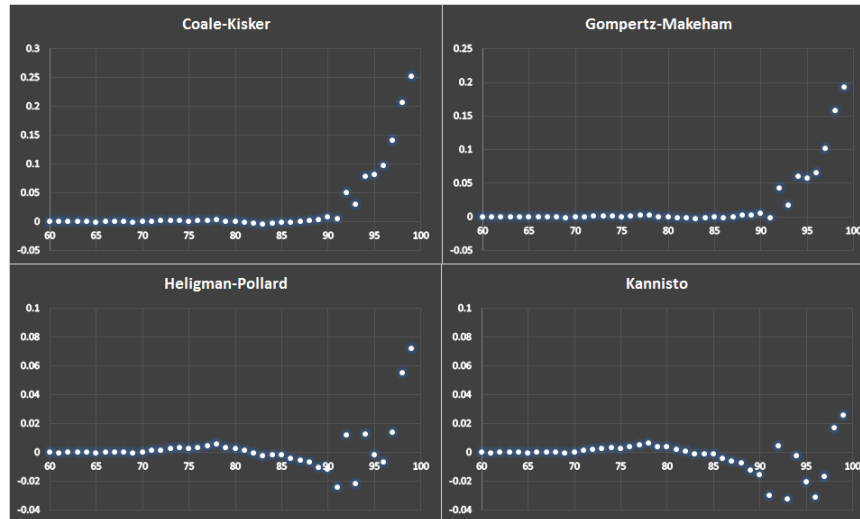
U žien je situácia podobná. Pripomeňme, že máme k dispozícii omnoho viac dát, výsledky u žien sú teda ďaleko spoľahlivejšie než tie u mužov. Obrázok 3.10 zachytáva vývoj MSEP za posledných 14 rokov u žien. V roku 2001 a 2002 boli všetky konkurujúce modely spoľahlivé. Od roku 2003 je už na prvý pohľad zrejmé, že Coale-Kiskerov, Gompertz-Makehamov a Gompertzov model sú spomedzi ostatných modelov najmenej presné. Za zmienku určite stojí pohľad na roky 2005 až 2010, kedy oficiálne používaný Gompertz-Makehamov model výrazne zaostával za modelmi Heligman-Pollarda, Kannista a Thatcher.



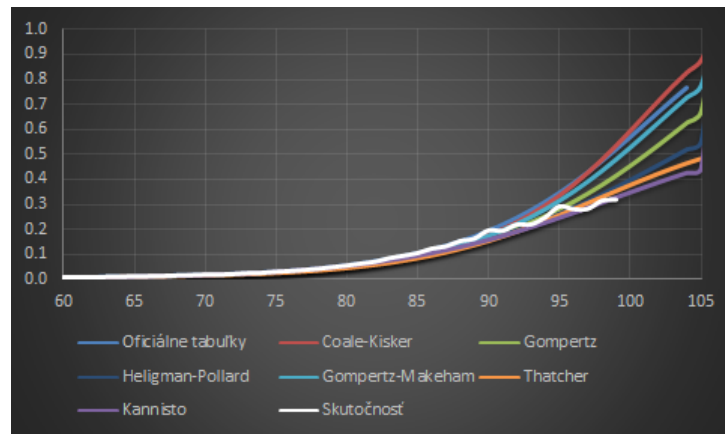
Obr. 4.10: Vývoj strednej štvorcovej chyby projekcie, ženy

Pri analýze reziduálnej zložky dospejeme k podobnému záveru. Modely Coale-Kiskera a Gompertz-Makehama často nadhodnocujú pravdepodobnosti úmrtia, najmä v rokoch najvyšších, na obrázku 3.11 je zachytená napríklad situácia z roku 2011. Všimnime si zmenu oboru hodnôt v grafoch u jednotlivých modelov, model Kannista sa javí byť priam ideálny pre skúmaný rok.

Vývoj úmrtnosti u žien potvrdzuje presvedčenia expertov, že úmrtnosť v Českej republike sa postupom času mení, tak ako sme to mohli v minulosti pozorovať v najrozvinutejších štátoch Európy, kde sa momentálne používajú logistické modely pre projekciu úmrtnosti v rokoch najvyšších. Môžeme tak očakávať, že do budúcnosti budú pravdepodobnosti úmrtia naďalej klesať a postupom času sa hádame dočkáme aj zmeny v oficiálnej metodológii tvorby úmrtnostných tabuliek s uprednostnením logistických modelov, tak ako je to napr. zrejmé na obrázku 3.12, ktorý zachytáva situáciu z roku 2010.



Obr. 4.11: Reziduá, ženy, 2011



Obr. 4.12: Pravdepodobnosti úmrtia, ženy, 2010

Vzhľadom na vývoj úmrtnosti sme sa v práci zaoberali aj voľbou maximálneho veku, ktorý ovplyvňuje výsledný tvar modelov vzhľadom na požadovanú konvergenciu pravdepodobnosti úmrtia v tomto veku. Jeho posunutie na hranicu 110, 115 alebo 120 ale neprinieslo žiadne vylepšenie v ktoromkoľvek modeli, preto v práci pracujeme s maximálnym vekom $\omega = 105$. V priloženom súbore nájdeme tvar úmrtnostných tabuliek aj pre tieto maximálne veku.

4.4 Predikčné intervaly pre projektované veku

Postup popísaný v sekcii 2.6.1 teraz aplikujeme na naše dáta, konkrétne sa budeme zaoberať modelom Heligman-Pollarda, pozri (3.17), pre ženy z roku 2014. Iteračnými algoritmami, popísanými v predchádzajúcej kapitole, sme získali odhady parametrov a a b , ktoré spolu so smerodajnou odchýlkou a 95 % intervalmi spoľahlivosti uvádza nasledujúca tabuľka:

	Odhad	Smerodajná odchýlka	Interval spoľahlivosti
a	0.117672	0.0127295	(0.0919023, 0.143441)
b	4.2301710^{-6}	5.0410^{-6}	$(-5.9727710^{-6}, 1.44331 \cdot 10^{-6})$

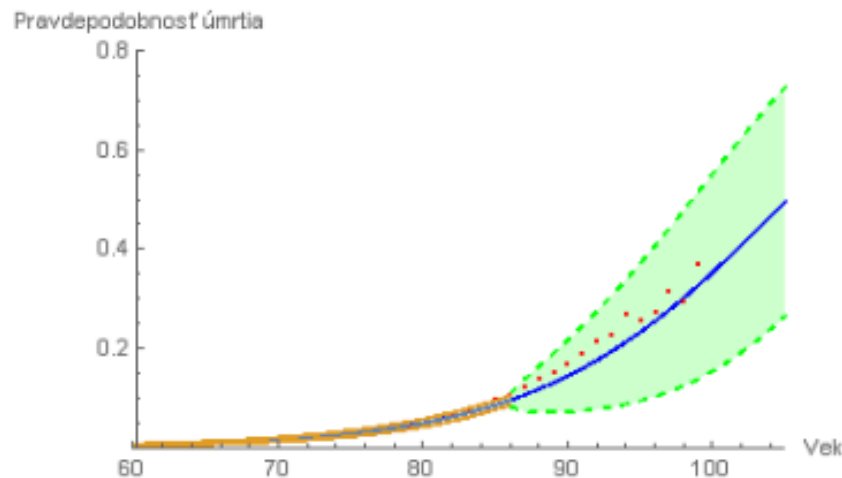
Kovariančná matica (3.33), získaná ako vedľajší produkt pri odhadovaní parametrov, má tvar

$$\Sigma = \begin{bmatrix} 0.000016204 & -6.4054410^{-9} \\ -6.4054410^{-9} & 2.5401610^{-12} \end{bmatrix}. \quad (4.4)$$

Rovnako tak získame aj podmienený rozptyl chybovej zložky, ktorý má v tomto prípade hodnotu 0.000171193. S využitím Taylorovej aproximácie dostávame odhad rozptylu rastovej funkcie pre $\tau = 1, \dots, 20$, pozri (3.35). Výsledný interval spoľahlivosti je spočiatku výrazne ovplyvnený aditívnou zložkou v podobe podmieneného rozptylu chybovej zložky, ktorý ale postupne stráca na dôležitosti a v rokoch najvyšších je jeho význam minimálny. Odhady týchto rozptylov sú zaznamenané v nasledujúcej tabuľke: (značíme $\text{var}(f_{x+\tau}(a, b, c)|r_x, \dots, r_0) = \text{VAR}$)

τ	VAR	τ	VAR
1	1.812 10^{-4}	11	6.390 10^{-3}
2	4.171 10^{-4}	12	7.496 10^{-3}
3	7.190 10^{-4}	13	8.640 10^{-3}
4	1.097 10^{-3}	14	9.791 10^{-3}
5	1.560 10^{-3}	15	0.0109101
6	2.116 10^{-3}	16	0.0119557
7	2.772 10^{-3}	17	0.0128855
8	3.531 10^{-3}	18	0.0136591
9	4.392 10^{-3}	19	0.0142414
10	5.349 10^{-3}	20	0.0146056

Teraz už môžeme zostrojiť výsledný 95% interval spoľahlivosti pre predpokladané pravdepodobnosti úmrtia v rokoch najvyšších (zelená). Ten je spolu so štandardným normálnym regresným intervalom pre veku 60 až 85 (hnedá) a skutočnosťou (červená) zachytený na obrázku 3.11. Môžeme si všimnúť, že 95% interval spoľahlivosti zachytil skutočne napozorované hodnoty, aj keď model pravdepodobnosti úmrtia v rokoch najvyšších v tomto roku mierne podhodnocuje. Zostrojené 95 % intervaly spoľahlivosti zachytili skutočné pravdepodobnosti úmrtia v absolútnej väčšine prípadov.



Obr. 4.13: Intervaly spoľahlivosti a skutočnosť

Daný postup môžeme použiť pre ľubovoľný konkurujúci model. 95 % intervaly spoľahlivosti pre projektované hodnoty majú podobné vlastnosti aj u ostatných modelov, na ich šírku spočiatku vplýva najmä podmienený rozptyl, jeho význam je ale do rokov najvyšších pri preferovaných modeloch relatívne malý v porovnaní s kovariančnou maticou (3.33). Čím spoľahlivejšia bola kvalita vyrovnaného modelu, tým sme získali užší interval spoľahlivosti.

Poznámka. V praktickej časti našej práce sme využívali hneď niekoľko špecializovaných softwarov. Odhady parametrov boli získané s použitím licenčného softwaru DERAS (DEathRatesSimulation), ktorý nám poskytla Prírodovedecká fakulta Univerity Karlovej, oddelenie demografie. Intervaly spoľahlivosti pre projektované veku boli napočítané ručne, s využitím licenčného softwaru Wolfram Mathematica, ktorý nám poskytla Matematicko-fyzikálna fakulta Univerzity Karlovej. Pre spracovanie a interpretáciu výsledkov sme využili Microsoft Excel 2010.

Kapitola 5

Vplyv na implicitnú hodnotu poisťovne

V nasledujúcej kapitole sa zameriame na dopad úmrtnostných tabuliek na poisťovňu. Predstavíme si pojem *implicitnej hodnoty* poisťovne v súlade s najaktuálnejšími smernicami, ktoré sú dostupné na webových stránkach CFO Forum v podobe [18] a [19]. Ako sme sa mohli presvedčiť v predchádzajúcej kapitole, oficiálne úmrtnostné tabuľky majú v posledných rokoch tendenciu nadceňovať pravdepodobnosti úmrtia, čo môže výrazne ovplyvniť riziko dlhovekosti.

5.1 Implicitná hodnota poisťovne

Implicitná hodnota poisťovne je ukazovateľ, ktorý vyjadruje hodnotu poisťovne z hľadiska akcionára. Existuje niekoľko verzií pre túto hodnotu, v našej práci sa budeme venovať tzv. tržne konzistentnej implicitnej hodnote alebo MCEV (Market Consistent Embedded Value). Ide o meradlo konsolidovanej hodnoty akcionárových záujmov v poisťovni krytom obchode, do ktorého spadajú aspoň poisťné zmluvy, ktoré sú regulátorom považované za dlhodobé životné. Táto hodnota je počítaná zo vzťahu

$$\text{MCEV} = \text{NAV} + \text{VIF}, \quad (5.1)$$

kde NAV (Net Asset Value) je čistá hodnota aktív a VIF (Value of in-force) je súčasná hodnota obchodu plynúca z poisťných zmlúv. Čistú hodnotu aktív pritom možno získať ako

$$\text{NAV} = \text{FS} + \text{RC}, \quad (5.2)$$

kde FS (Free Surplus) sú voľné aktíva a RC (Required Capital) je požadovaný, viazaný kapitál. Súčasnú hodnotu obchodu napočítavame zo vzťahu

$$\text{VIF} = \text{PVFP} - \text{TVFOG} - \text{FCRC} - \text{CRNHR}, \quad (5.3)$$

kde PVFP (Present Value of Future Profits) je súčasná hodnota budúcich ziskov, ktorú je na základe dostupných predpokladov nutné odhadnúť. Úmrtnostné tabuľky vstupujú do výpočtov všetkých ukazovateľov v (5.3), ale vplyv na ostatné ukazovatele je v porovnaní s vplyvom na PVFP zanedbateľný. TVFOG (Time Value of Options and Guarantees) je časová hodnota opcí a garancií, FCRC (Fictional Costs on Required Capital) sú frikčné náklady na viazaný kapitál a CRNHR

(Cost of Residual Non Hedgable Risks) sú náklady na zostávajúce nezaistené riziká.

Poznámka. K výpočtu MCEV sa používajú okrem iného nasledujúce vstupy a predpoklady: *informácie o poistných zmluvách*, ktoré sa získavajú z prevádzkového systému a často sa zoskupujú do tzv. modelpointov, čo sú zväzky poistných zmlúv s rovnakými parametrami, napr. frekvencia platenia poistného, pohlavie či produkt. Informácie o *aktivách poisťovne*, napr. nominálna hodnota, kupón, splatnosť dlhopisu, akcie alebo investičného fondu a tzv. *výpočetné podklady druhého rádu*, do ktorých spadajú najlepšie odhady ekonomických, biometrických a pre danú poisťovňu špecifických ukazovateľov, veličín. Jedným z hlavných biometrických ukazovateľov je práve úmrtnosť a jej správanie.

Pozrime sa teraz na prognózu nákladov a výnosov v súvislosti s celkovou očakávanou poistnou čiastkou v jednoduchom cash flow modeli zmiešaného poistenia, bez priznaného podielu na výnosoch. V ďalších výpočtoch budeme pre jednoduchosť používať nasledujúce označenia:

PC(0)...celková poistná čiastka všetkých poistných zmlúv a
 PC(m)...očakávaná hodnota poistnej čiastky po m rokoch trvania poistenia,
 v^* ...podiel rezervy poistného vrátený pri storne,
 i^* ...najlepší odhad technickej úrokovej miery,
 s_{kn}^* ...najlepší odhad očakávaného podielu storien v k -tom roku poistenia, zjednané na n rokov.

Použijeme takisto početné podklady 2. rádu, pod ktorými rozumieme najlepšie odhady (best estimate) týchto parametrov:

q_x^* ...pravdepodobnosť, že osoba vo veku x sa nedožije veku $x+1$,
 α^* ...náklady na sprostredkovanie poistenia, ako % z poistnej čiastky,
 β^* ...bežné (ročné) správne náklady v % z poistnej čiastky,
 γ^* ...inkasné, ročné náklady v % z poistnej čiastky.

Pre najlepší odhad vývoja poistnej čiastky potom platí indukčný vzťah

$$PC^*(m) = (1 - q_{x+m-1}^*)(1 - s_{mn}^*)PC^*(m-1). \quad (5.4)$$

Očakávaný zisk v roku $m+1$ EI (expected income) potom možo určiť z počiatočnej hodnoty budúcich peňažných tokov v tomto roku, ku ktorej pripočítame poistné a odpočítame správne náklady, poistné plnenie, ľuďom, ktorí zomreli, storná, o ktorých predpokladáme, že sa odohrávajú na konci roka a hodnotu budúcich peňažných tokov, ktoré musíme mať k dispozícii, to jest

$$\begin{aligned} EI^*(m) = & V(m)PC^*(m)(1 + i^*) + P PC^*(m)(1 + i^*) \\ & - (\beta^* + \gamma^* P)PC(m)(1 + i^*) - q_{x+m}^*PC(m) \\ & - (1 - q_{x+m}^*)s_{m+1n}^*v^*V(m+1)PC^*(m) - V(m+1)PC^*(m+1), \end{aligned} \quad (5.5)$$

kde $V(m)$ je súčasná hodnota budúcich tokov na jednotku poistnej čiastky a P je výška stanoveného poistného. Správne odhadnuté pravdepodobnosti úmrtia sú teda potrebné pre dobré vyhodnotenie poistného plnenia a storien, čo priamo súvisí s výpočtom MCEV prostredníctvom PVFP.

5.2 Dopad na poisťovňu

Ako sme sa mohli presvedčiť v predchádzajúcej sekcii, úmrtnostné tabuľky sú dôležitým predpokladom pre výpočet očakávaného zisku v konkrétnom roku, PVFP a MCEV. Nesprávne určenie pravdepodobností úmrtia tak môže mať negatívny vplyv na distribúciu financií v rámci poisťovne. V prípade, že upravíme rovnicu (5.6) do tvaru

$$\begin{aligned} EI^*(m) = & PC^*(m)[V(m) + P(1 - \gamma) - \beta](i^* - i) \\ & + PC^*(m)(q_{x+m} - q_{x+m}^*)[1 - V(m + 1)] \\ & + (1 - q_{x+m}^*)s_{m+n}^*V(m + 1)[1 - v^*] \\ & + PC^*(m)(1 + i^*)(\beta - \beta^* + (\gamma - \gamma^*)P), \end{aligned} \quad (5.6)$$

dostávame rozklad výnosu na úrokový nadvýnos, výnos z úmrtnosti, výnos zo storien a výnos z úspory správnych nákladov. Výnos z úmrtnosti v tomto prípade realizujeme, ak ľudia umierajú viac, než sme očakávali, to jest, náš model pravdepodobnosti úmrtia nadhodnocoval. Je zrejmé, že v prípade, že by náš model pravdepodobnosti úmrtia výrazne podhodnocoval, mohli by sme zaznamenať nepríjemnú stratu a zbytočne sa vystavovať riziku nesolventnosti.

Nadhodnocovanie pravdepodobností úmrtia ale neznamená zisk pre poisťovňu. Ponuka poisťných produktov je v dnešnej dobe veľmi pestrá a ak uvažíme napr. poistenie pre prípad dožitia, je nadhodnocovanie pravdepodobností úmrtia veľmi nebezpečné. Preto je v najlepšom záujme poisťovne zachytiť tieto pravdepodobnosti čo najpresnejšie. Tvrdenie, ktoré sa v poisťovni často používa: *Aktuár je človek, ktorý chce, aby ste umreli presne načas*, je plne opodstatnené.

Príklad. Uvažujme modelový poisťný kmeň o veľkosti 100 000, ktorý tvoria výhradne ženy vo veku 25 až 60 rokov. Množstvo žien v jednotlivých rokoch bolo normalizované podľa oficiálnych údajov o počte žien k 1. 1. 2014, dostupné na stránkach ČSÚ. Uvažujme, že si tieto ženy v roku 2004 zjednali dôchodkové životné poistenie, ktoré im od veku 65 rokov začne vyplácať pravidelnú, ročnú annuitu vo výške 1 000 Kč, vždy na začiatku roka, pokiaľ je poistená osoba nažive. Naším cieľom bude odhadnúť diskontovanú hodnotu poisťných plnení spojenú s týmto poisťným kmeňom, pričom nás bude zaujímať jej hodnota v roku 2014. Výpočet najskôr prevedieme so spočítanou úmrtnostnou tabuľkou podľa modelu Gompertz-Makehama pre ženy na rok 2004. V tomto prípade teda nebudeme uvažovať vývoj úmrtnosti za roky 2005 až 2014. Úmrtnostné tabuľky z týchto rokov následne zahrnieme do druhého výpočtu a oba výsledky medzi sebou porovnáme. Rovnaký postup potom aplikujeme na model Kannista.

Ide o úlohu na životný predľhotný dôchodok, odložený o m rokov, kde m je počet rokov, ktoré chýbajú danej osobe k dovŕšeniu veku 65 v roku 2014. U žien v rokoch 60 až 64 ide o klasický predľhotný dôchodok, nakoľko im sa už poisťné plnenie začalo vyplácať, t. j. v tomto prípade $m = 0$. Nech K je náhodná premenná, ktorá značí počet rokov, ktoré nastanú od uzavretia poistenia až po smrť poistenej osoby. Diskontovaná hodnota poisťných plnení Y má tvar

$$\begin{aligned} Y = & 0, & K < m \\ = & v^m + \dots + v^K, & K \geq m. \end{aligned} \quad (5.7)$$

Strednú hodnotu poistných plnení môžeme spočítať podľa [3] napríklad s využitím komutačných čísel

$$\mathbf{E}Y = \frac{N_{x+m}}{D_x}. \quad (5.8)$$

Následne upravíme množstvo žien z roku 2004, ktoré sme stanovili hore uvedeným spôsobom, podľa úmrtnostných tabuliek z roku 2004 v prvom prípade, a z rokov 2004 až 2014 v druhom prípade. Rovnako tak u modelu Kannista. Hodnoty komutačných čísel vychádzajú z pravdepodobností úmrtí, ktoré sú pri týchto dvoch prístupoch, ale aj v rámci jednotlivých modelov rôzne od veku 60. Tým pádom dostávame rôzne očakávané diskontné hodnoty poistných plnení, čím sa mení aj hodnota PVFP. Použijeme technickú úrokovú mieru 1.5%. V priloženom súbore nájdeme možnosť nastaviť veľkosť poistného kmeňa, hodnotu ročnej výplaty a technickú úrokovú mieru, výpočet sa automaticky sám aktualizuje. Celková hodnota diskontovaných poistných plnení je rozložená podľa vstupného veku vedľa úmrtnostných tabuliek.

V prvom prípade sme neuvažovali vývoj úmrtnosti od roku 2004 a hodnotu diskontovaných poistných plnení, vzťahnutú k roku 2014, sme počítali len z úmrtnostných tabuliek z roku 2004. Obdržané výsledky zachytáva tabuľka MCEV, vzhľadom na vstupné veku 56 až 60 uvádzame aj čiastku, ktorú sme už zaplatili.

Model	Celkovo zaplatíme	Zaplatili sme	Zostáva zaplatiť
Gompertz-Makeham	1 152 552 733	38 781 921	1 113 770 813
Kannisto	1 151 914 818	38 790 747	1 113 124 071

Obr. 5.1: Tabuľka MCEV: diskontovaná hodnota poistných plnení

Na prvý pohľad je prekvapivé, že v prípade Kannistovho modelu očakávame, že celkovo zaplatíme menej ako u Gompertz-Makehama, napriek tomu, že pravdepodobnosti úmrtia sú u tohto modelu v rokoch najvyšších výrazne nižšie. Úmrtnostné tabuľky sme ale vyrovnávali už od veku 60, Kannistov model bol až do veku 85 trochu viac skeptický a vyrovnané pravdepodobnosti úmrtia boli o niečo málo vyššie, čo vysvetľuje získané výsledky.

V druhom prípade sme uvažovali vývoj úmrtnosti od roku 2004 a hodnotu diskontovaných poistných plnení, vzťahnutú k roku 2014, sme počítali na základe informácií zo všetkých úmrtnostných tabuliek za dané obdobie rokov 2004 až 2014. Obdržané výsledky zachytáva tabuľka MCEV2, rovnako ako v prvom prípade uvádzame aj hodnotu, ktorú sme už zaplatili.

Model	Celkovo zaplatím	Zaplatili sme	Zostáva zaplatiť
Gompertz-Makeham	1 280 007 623	38 940 264	1 241 067 359
Kannisto	1 275 104 713	38 930 543	1 236 174 170

Obr. 5.2: Tabuľka MCEV2: diskontovaná hodnota poistných plnení

Vidíme, že diskontovaná hodnota poistných plnení je výrazne vyššia než tomu bolo v prvom prípade. Ako sme už spomínali vyššie, pravdepodobnosti úmrtia pre

veky najvyššie za toto obdobie výrazne poklesli. To znamená, že poistené ženy sa budú dožívať vyšších rokov a my tak budeme vyplácať zjednanú čiastku dlhšie. Za 10 rokov to činí rozdiel približne 125 miliónov, ktorá zodpovedá negatívnemu dopadu na PVFP za sledované obdobie. Celkovo sme za 10 rokov zaznamenali zhruba 14.75 % nárast, čo odpovedá nárastu 1.31% za jeden rok. Δ

Nemenej dôležitou súčasťou výpočtu MCEV je aj analýza pohybu oproti minulému stavu. Zmena MCEV môže byť vyvolaná zmenou spôsobenou zmenou modelu, zmenou spôsobenou zmenou v čase, zmenou portfólia, vplyvom nového obchodu, zmenou vplyvom zmeny predpokladu. Táto analýza býva pritom súčasťou oficiálneho reportu. To znamená, že akákoľvek zmena, napr. úmrtnostných tabuliek, musí byť v reporte uvedená s presným dopadom na pohyb MCEV.

Vplyv úmrtnostných tabuliek je možné nájsť aj v analýze citlivosti MCEV na konkrétne predpoklady. CFO princípy [18] a [19] predpisujú testovať citlivosť v minimálnom rozsahu:

- zníženie výnosovej krivky na trhu o 100 bazických bodov (1%)
- nárast rizikovej diskontnej miery o 100 bazických bodov (1%)
- 10 % zníženie kapitálovej hodnoty akcií/nehnutelnosti k dátumu ocenenia
- 25 % zvýšenie v majetkovo implikovanej volatilitate k dátumu ocenenia
- 10 % zníženie správnych nákladov
- 10 % proporčné zníženie stornovosti
- 5 % proporčné zníženie úmrtnosti (zvlášť pre poistenie životné a annuitné)

Požadovaný kapitál by mal byť rovný levelu solventnostného kapitálu.

Tieto testy slúžia najmä k identifikácii rizikových faktorov pre portfólio a poskytujú návod ako ho riadiť. V praxi sa citlivostná analýza testuje jednak pre celé portfólio, a jednak pre nový obchod. Predpísaná perióda testovania je minimálne raz do roka, veľa poisťovní ale vyžaduje zverejnenie oficiálnych výsledkov z citlivostnej analýzy pri každom kvartálnom reporte.

Kapitola 6

Záver

V práci sme sa venovali najpoužívanejším demografickým modelom pre projekciu do rokov najvyšších v súvislosti s tvorbou úmrtnostných tabuliek a ich vplyvu na implicitnú hodnotu poisťovne. Na základe množstva a kvality dát sme určili hornú vekovú hranicu pre odhadovanie parametrov na vek 85 rokov, od ktorého začíname s projekciou.

V praktickej časti sme najskôr porovnali kvalitu vyrovnaní jednotlivých modelov pre roky 60 až 85 na populácii Českej republiky. Za posledných 15 rokov, ktoré boli našim primárnym cieľom, sme dospeli k veľmi podobným výsledkom, o žiadnom modeli sa nedá prehlásiť, že by bol pre účely vyrovnaní nevhodný pre toto vekové rozpätie, aj keď Coale-Kiskerov model podával najmenej uspokojivé výsledky u oboch pohlaví. Je potrebné si ale uvedomiť, že tento model bol navrhnutý až pre vyššie veku, preto by nás jeho relatívne horšie vyrovnanie nemalo nejako výraznejšie prekvapiť.

Pri projekcii do rokov najvyšších sme už ale obdržali výrazné rozdiely. U mužov, kde bola analýza náročnejšia vzhľadom na menší počet a väčšiu volatilitu dát, sa do veku 90 väčšinou javili modely Gompertz-Makehama a Coale-Kiskera ako najvhodnejšie. Od veku 90 sme ale u týchto modelov zaznamenali značné nadhodnocovanie pravdepodobností úmrtia, najmä v rokoch 2006 až 2011, kedy by sme uprednostnili logistické modely alebo model Heligman-Pollarda. V posledných 3 rokoch si boli modely navzájom dobre konkurujúce. Vzhľadom na väčšiu volatilitu a občasné rozdielne správanie sa úmrtnosti v po sebe nasledujúcich rokoch by sme nakoniec asi uprednostnili model Heligman-Pollarda, tesne pred logistickými modelmi.

U žien sú výsledky ešte viac výrazné. Modely Coale-Kiskera a Gompertz-Makehama v sledovanom období posledných 15 rokov jednoznačne nadhodnocujú pravdepodobnosti úmrtia pre roky najvyššie. Naopak logistickým modelom Thatcher a Kannista sa pravdepodobnosti úmrtia v týchto vekoch podarilo vysvetliť najlepšie, predovšetkým Kannistov model bol v období rokov 2006 až 2011 mimoriadne presný. Okolo veku 90 sme ale aj tu občas napozorovali mierne podhodnotenie. Tento model sa naviac využíva pri projektovaní rokov najvyšších v rámci projektu Human Mortality Database a je aj oficiálnym modelom pre konštrukciu úmrtnostných tabuliek napr. vo Švédsku. V budúcnosti by sa preto mohol stať oficiálnym modelom aj u nás.

Nami získané výsledky tak potvrdili kritiku expertov, že Gompertz-Makehamov model nie je pre roky najvyššie úplne najvhodnejší, keďže tu pravdepodobnosti úmrtia zrejme nadhodnocuje. Vzhľadom na paralelu vývoja úmrtnosti medzi jednotlivými štátmi vyspelého sveta, či paralelu medzi ženami a mužmi môžeme očakávať, že postupom času sa logistické modely stanú jednoznačne najlepšími kandidátmi na projekciu úmrtnosti do rokov najvyšších pre obe pohlavia. Veríme, že sa v priebehu niekoľkých rokov dočkáme aj zmeny oficiálnej metodológie pre tvorbu úmrtnostných tabuliek v Českej republike.

Literatúra

- [1] *Burcin, B., Tesárková, K., Šidlo, L.*: “Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci.” *Demografie* 52, 2010: 77-89
- [2] *Carriere, F.J.*: Parametric models for life tables. *Transactions of Society of Actuaries*, Volume 44, 1992.
- [3] *Cipra, T.*: Pojistná matematika: teorie a praxe, 2. aktualizované vydání. Ekopress, Praha, 2006
- [4] *Dennis, J.E., Jr.*: Nonlinear least squares and equations. Academic Press, London, 1977.
- [5] *Dr. Buettner, T.*: Approaches and Experiences in Projecting Mortality Patterns for the Oldest-Old. *North American Actuarial Journal*, Volume 6, Issue 3, 2002.
- [6] *Gavrilov, L.A., Gavrilova, N.S.* : The biology of Life Span: A Quantitative Approach. Harwood Academic Publishers, Švajčiarsko, 1991.
- [7] *Gompertz, B.*: On One Uniform Law of Mortality from Birth to Extreme Old Age, and on the Law of Sickness. *Journal of the Institute of Actuaries*, 16:3229-44, 1872.
- [8] *Horiuchi, S., Coale, A.J.*: Age Patterns of Mortality for Older Woman: An analysis using Age-Specific Rate of Mortality Change with Age. *Mathematical Population Studies* 2(4):245-67, 1960.
- [9] *Horiuchi, S., Wilmoth, J.*: Deceleration in the Age Pattern of Mortality at Older Ages. *Laboratory of Populations*, Rockefeller University, 1996
- [10] *Meade, N., Islam, T.*: Prediction intervals for growth curve forecasts. John Wiley & Sons, Inc., London, 1995.
- [11] *Mildvan, A., Strehler, B.L.*: A Critique of Theories of Mortality. *Biology of Aging*, Washington, 1960.
- [12] *Olshansky, J.S., Carnes, A.B.*: Ever since Gompertz. *Demography*, Volume 34, Issue 1, pp 1-15, 1997.

- [13] *Pavlík, Z., Rychtaříková, J., Šubrtová, A.*: Základy demografie. Academia, Praha, 1986.
- [14] *Seber, G.A.F, Lee, J.A.*: Linear Regression Analysis. Vol.936. John Wiley & Sons, 2012.
- [15] *Seber, G.A.F., Wild, C.J.*: Nonlinear Regression. John Wiley & Sons, Inc., New Jersey, 2003.
- [16] *Stuart, A., Ord, K.J.*: Kendall's Advanced Theory of Statistics: Distribution Theory. Griffin, London, 1987.
- [17] *Thatcher, R.A.*: The long-term pattern of adult mortality and the highest attained age. Journal of the Royal Statistical Society: Series A, Volume 162, Issue 1, 1999.
- [18] MCEV Basis for Conclusions.pdf, April 2016, priebežne aktualizované a dostupné na http://www.cfoforum.nl/embedded_value.html
- [19] MCEV Principles & Guidelines.pdf, April 2016, priebežne aktualizované a dostupné na http://www.cfoforum.nl/embedded_value.html

Zoznam obrázkov

4.1	Pravdepodobnosti úmrtia v rokoch 1960 a 1961, muži	16
4.2	Pravdepodobnosti úmrtia v rokoch 1974 a 1984, ženy	17
4.3	Odhady parametrov v jednotlivých kalendárnych rokoch	18
4.4	Vekovo špecifická miera úmrtnosti	18
4.5	Tabuľka 1: MSE, muži, 2008	19
4.6	Tabuľka 2: MSE, ženy, 2008	20
4.7	Vývoj strednej štvorcovej chyby projekcie, muži	21
4.8	Reziduá, muži, 2010	21
4.9	Pravdepodobnosti úmrtia, muži, 2009	22
4.10	Vývoj strednej štvorcovej chyby projekcie, ženy	22
4.11	Reziduá, ženy, 2011	23
4.12	Pravdepodobnosti úmrtia, ženy, 2010	23
4.13	Intervaly spoľahlivosti a skutočnosť	24
5.1	Tabuľka MCEV: diskontovaná hodnota poistných plnení	29
5.2	Tabuľka MCEV2: diskontovaná hodnota poistných plnení	29